Raytheon

Air Traffic Management System Development and Integration (ATMSDI)

CTO-02 – Human Factors Support for DAG-TM Simulation

CTOD 2.10-3 – Final Report

May 2003

Contract No. NAS2-00015

Prepared for

NASA Ames Research Center
Moffett Field, CA 94035-1000

Prepared by

Raytheon ATMSDI Team
1001 Boston Post Road
Marlborough, MA 01752-3789

EXECUTIVE SUMMARY

Background

The Distributed Air-Ground Traffic Management (DAG-TM) concept represents a possible approach to the implementation of the free flight concept defined by the RTCA (1995). DAG-TM is a National Aeronautics and Space Administration (NASA) effort led by three research centers (i.e., Ames, Langley, and Glenn) and funded by the Aviation Systems Program under the Advanced Air Transportation Technologies (AATT) project. DAG-TM describes operational modes that would increase the flexibility, efficiency, and capacity of the National Airspace System (NAS) by providing stakeholders (i.e., pilots, air traffic service providers [ATSP], and Airline Operations Center [AOC] specialists) with decision support tools (DST) that promote collaborative decision making and distributed problem solving. The DAG-TM concept definition (AATT, 1999) includes 15 conceptual elements that address all user classes and cover all flight phases (i.e., pre-flight planning, departure, cruise, and arrival) and operational domains in the NAS (i.e., surface, en route airspace, and terminal airspace). The goal of DAG-TM is to provide these enhancements without adversely affecting system safety or restricting user accessibility to the NAS.

Defining characteristics of DAG-TM concept elements include the redistribution of roles and responsibilities for maneuvering and separation assurance from the ATSPs to the flight deck (FD). Operating in a distributed and cooperative problem-solving environment will affect stakeholder information needs and change traditional roles and responsibilities. Fundamental human factor issues center on a redistribution of workload and a need for shared situation awareness to foster safe and effective communication and coordination.

NASA Ames Research Center activities to date have included several workshops (May 2001, August 2001, January 2002) and demonstrations (September 2001 and June 2002). These activities have centered on the development and testing, via user feedback, of the advanced technologies necessary for operating in a distributed environment. These activities also have been an avenue for refining the concept definitions within the scope of DAG-TM.

The DAG-TM Concept Elements

DAG-TM research conducted at the NASA Ames Research Center has specifically focused on the development and evaluation of three of the fourteen concept elements: en route trajectory negotiation (CE 6), en route free maneuvering (CE 5), terminal approach self-spacing (CE 11).

En Route Trajectory Negotiation involves real-time collaboration among the DAG-TM stakeholders for implementing user-preferred trajectory changes. The concept element proposes to enhance trajectory negotiation by: increasing ATSP trajectory prediction capability through user-supplied data on key flight parameters, and by supplying the user (i.e., FD or AOC) with the capability to formulate conflict-free trajectory changes that will conform to active traffic flow management (TFM) constraints (if equipped). In this mode of operation, ATSPs are still responsible for separation and still need to approve trajectory change requests.

I In *En Route Free Maneuvering*, appropriately equipped aircraft accept the responsibility to maintain separation from other aircraft while exercising the authority to freely maneuver in en route airspace. Free maneuvering aircraft have the authority to establish new user-preferred trajectories with the restriction that new trajectories conform to local TFM constraints and do not create traffic conflicts.

Terminal Approach Self-Spacing is intended to increase terminal area throughput by providing flight crews and ATSPs with a method for reducing in-trail separation. Appropriately equipped aircraft entering an arrival stream are assigned an in-trail spacing interval behind a lead aircraft. The time-based spacing interval, rather than distance-based, allows for spacing compression as aircraft speeds decrease throughout the approach.

September 2002 Simulation

In September 2002, the DAG-TM research team conducted a large-scale, high fidelity human-in-the-loop simulation to investigate the feasibility of the three concept elements. The study was conducted at the NASA Ames Research Center and included one off-site remote flight deck station located at San Jose State University.

Because most of the previous DAG-TM work had consisted of demonstrations and workshops, the main objective of the September 2002 simulation was to design and conduct a study that would permit the statistical analysis of operational and subject ratings data. In addition, there were several specific goals of the simulation. The first was to form an initial assessment of the operational viability and potential benefits of the concept elements by comparing DAG-TM operational conditions with a baseline environment that represented current-day technologies. The second goal was to evaluate the tools and the changing responsibilities, tasks, and interactions between pilots and ATSPs. Results regarding pilot and ATSP ability to function in a manner prescribed by the concept elements were used to further define these concepts. Last, the goal was to assess the current status of the simulation technology to support DAG-TM research and identify areas of development for future studies.

Experimental Design

Laboratory facilities were configured to simulate the airspace that handles the northwest arrivals into the Dallas/Forth Worth area. The airspace encompassed en route sectors from three centers (Albuquerque, Fort Worth, and Kansas City), the Dallas/Forth Worth Terminal Radar Approach Control (TRACON), and two airports (Dallas/Fort Worth International Airport and Dallas Love Field). Traffic scenarios were about 1-hour long and comprised approximately 90 aircraft. Five full performance level controllers staffed four en route sectors and one TRACON sector. Three confederate controllers staffed two peripheral en route sectors and one TRACON sector.

Aircraft equipage was categorized as either equipped or unequipped. Unequipped aircraft were minimally equipped with a flight management system (FMS), cockpit situation display (CSD) with Traffic alert and Collision Avoidance System (TCAS), required time of arrival (RTA) capability, a time-based arrival self-spacing tool, automatic dependent surveillance-broadcast (ADS-B), and controller pilot data link communications (CPDLC). Equipped aircraft, in addition, possessed a CSD with aircraft intent information, conflict detection tool with manual resolution (CD&R), and a route-planning tool. Eight air transport rated pilots participated in the study. Six pilots flew the equipped aircraft via PCPlane workstations while two other pilots teamed up (first and second officer) to fly the Advanced Concepts Flight Simulator (ACFS), a high fidelity, full-mission flight simulator. Pseudo-pilots flew the remaining aircraft, all unequipped, via Multi Aircraft Control System (MACS) workstations.

The study included 12 simulation runs conducted under one of three experimental conditions (or modes of operation):

Baseline Designed to represent present-day operations with the addition of TMA.

CE 6 In the trajectory negotiation mode of operation, equipped aircraft were allowed to make trajectory change requests that could be data-linked to the controller for approval. The TRACON controller was able to issue self-spacing clearances to all aircraft.

CE 5 Equipped aircraft in the free maneuvering mode (i.e., autonomous aircraft) had the authority to initiate flight path changes in en route airspace without ATC approval. Autonomous aircraft were responsible for their separation while controllers remained responsible for the separation of managed aircraft (aircraft not in free maneuvering mode). Separation of conflicting aircraft was dictated by a set of rules-of-the-road. Terminal approach self-spacing was possible for all aircraft.

Benefits in Efficiency and Predictability

The proposed benefit of CE 6 and CE 5 is that the flight crew will have more autonomy in choosing their flight path and, therefore, will be able to fly more efficient routes. With the advanced technologies available, equipped aircraft can stay on a preferred trajectory longer by using the vertical navigation (VNAV) descent and speed advisories to meet RTAs instead of using vectors. The addition of improved CD&R and route planning capabilities will enable flight path modifications that reduce non-preferred deviations.

Several metrics were used to measure system efficiency and predictability in the en route environment, which encompassed en route operations from approximately 200 nautical miles outside the TRACON boundary to the meter fix and included the initial descent phase of flight. Analyses of the system-level metrics revealed some significant differences between conditions. No multiple comparison tests could be run to determine between which conditions the differences were significant; but the descriptive analyses of arrival delivery accuracy, altitude deviation, en route altitude, distance traveled, and flight time all suggest that runs in the Baseline condition were less efficient and/or predictable.

controllers and pilots of autonomous aircraft could maintain the arrival schedule. The assumption is that a more predictable feed to the TRACON would reduce the need for spacing buffers and, therefore, increase the rate at which aircraft could be accepted into terminal airspace. Absolute arrival delivery error varied significantly across conditions. En route controllers were able to provide a steady and efficient feed to the TRACON controller in the CE 6 and CE 5 conditions, while in the Baseline, the feed was less representative of the actual arrival schedule, with an average absolute arrival delivery error three times greater.

It should be noted that a disproportionate number of resequenced flights were not entered using the swap function during the Baseline runs. The failure to use the swap function precluded the resequenced flights from having their scheduled time of arrival (STA) updated, which consequently inflated the absolute delivery error. It has not yet been determined why controllers tended not to use the swap function in the Baseline condition.

Altitude Deviation. There were no significant differences between conditions for average altitude deviation for flights off-altitude. However, the proportion of flights that were off-altitude was significantly larger in the Baseline condition than in the CE 6 and CE 5 conditions. It appears that aircraft were as good, if not better, at meeting the meter fix crossing restriction when flying the FMS-generated precision descent (CE 6 and CE 5 conditions) as when relying on controller clearances (Baseline condition).

En Route Altitude. The average en route altitude at 60 nautical miles from the meter fix, which constitutes the initial descent phase, varied significantly by condition, with the lowest in the Baseline condition. The difference in average altitude is likely a result of the different descent procedures. The precision descent procedure permitted pilots to fly the aircrafts' FMS descent profile, which generally allowed aircraft to maintain cruise altitude longer and start the descent later in the flight.

Distance Traveled and Flight Time. Results showed that flights in the Baseline condition flew significantly more miles than in the CE 6 and CE 5 conditions. Controllers did not have speed advisories available in the Baseline condition, which may have led to more vectoring, as opposed to speed changes, when aircraft needed to absorb delay. Similarly, flight time varied significantly across conditions, with average flight times in the CE 6 and CE 5 conditions less than in the Baseline condition. In addition to the extra miles flown, the increase in flight time in the Baseline condition could have been due to different descent procedures employed. The Baseline condition used a standard descent procedure that resulted in aircraft being descended earlier and stepped down throughout the descent, which slowed an aircraft's initial descent to the meter fix as compared to the precision descent procedure employed in the CE 6 and CE 5 conditions.

Operational Viability

An important stipulation of the DAG-TM concept is that gains in predictability, flexibility, and efficiency should not compromise safety. Two system-level metrics used during the simulation, separation violations and arrival spacing, provided a measure of overall safety.

Separation Violations. Results suggest CE 5 and CE 6 operations did not decrease overall safety, as indicated by fewer violations occurring in the CE conditions than in the Baseline. In fact, safety in the en route sectors may have been increased, possibly due to the added functionalities of the ground- and airside tools. Overall, only one violation occurred near the meter fix where the heaviest congestion occurred and the only separation violations involving PCPlane flights occurred in the Baseline condition.

Arrival Spacing. Arrival spacing was a measure of the extent to which controllers and pilots deviated from a 58-second separation minimum. Results showed that the proportion of flights and their average deviation did not vary significantly between conditions. These findings suggest that controllers had no more difficulty in maintaining the minimum spacing requirement during the CE 6 and CE 5 runs than during the Baseline runs.

A main premise of the DAG-TM concept is that an air-ground distributed control system will better accommodate traffic growth. However, the viability of the concept is dependent on maintaining an acceptable workload level in light of shifting roles and responsibilities. Generally, controller and pilot subjective workload ratings showed that workload was not adversely affected by the changes in roles and responsibilities.

Controller Workload. Differences between the average workload ratings (using the Air Traffic Workload Input Technique) in the three conditions were significant for only one high altitude position, but all average workload ratings were highest in the Baseline condition. En route controllers also stated in post-run questionnaires that their workload, mental demand, effort, and frustration were higher, on average, in the Baseline than in the CE 6 and CE 5 conditions. The opinion of the TRACON controller contrasted with those of the en route controllers because the controller's perceived workload was higher during terminal approach spacing runs than during Baseline runs.

Pilot Workload. Descriptive analyses of the subjective judgments of workloads made by pilots after each run and at the end of the study suggest that perceived workload did not vary between experimental conditions. It was perceived to be comparable to, or slightly lower than, the level experienced during normal current-day operations. One reason why subjective ratings of workload seem not to have varied between experimental conditions may be that the level of complexity or difficulty of the scenarios was not high enough for the pilots to fully need the support provided by the advanced technologies.

Acceptability of the Concept Elements

Participants rated the acceptability of the three concept elements with respect to normal current-day operations. Pilots favorably rated the three concept elements while; controllers favorably rated only trajectory negotiation (CE 6). Controllers felt that free maneuvering (CE 5) was less acceptable when compared to normal current-day operations. Most concerns related to procedural issues in dealing with traffic conflicts and managing aircraft off their STAs. Similarly, the TRACON controller found terminal spacing (CE 11) to be less acceptable than normal current-day operations, with the biggest concern being when to initiate in-trail spacing.

Usability of DAG-TM Technologies

Pilots evaluated the usefulness of the CSD features and generally provided positive feedback, but average ratings and their dispersion suggest that improvements could be made. Pilots found that data link helped reduce workload in the en route environment. Most CSD features received a positive usability rating. Again, average ratings were often only slightly higher than the neutral point; and some variation of opinion among pilots suggests that the usability of some features could be improved.

En route controllers' ratings of the adequacy of their computer interface were slightly higher than the neutral point and similar across conditions. Ratings of individual interface features showed that most items were considered useful and usable. Only two features, the conflict list and the route modification tool, received neutral average ratings. Controllers found the interface very adequately distinguished between autonomous and managed aircraft. The TRACON controller evaluated the usability and usefulness of some CE 11 displays and tools. All items received positive ratings, except the shortcut window for handoffs and runway change information were judged very difficult to use and unnecessary.

Conclusion

The DAG-TM research team successfully implemented an advanced-concept simulation environment that demonstrated the benefits and feasibility of an air traffic management system involving distributed, although limited, maneuvering and separation responsibilities. Controllers and pilots were able to accomplish their assigned tasks while showing increases in system efficiency under experimental conditions. The results of the study indicate that the CEs have the potential to improve system efficiency and that they were well received by most participants. The study also provided valuable information regarding potential improvements of the DSTs, procedures, and research methodology.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

1. INTRODUCTION

In September 2002, a research team at the National Aeronautics and Space Administration (NASA) Ames Research Center (ARC) conducted a human-in-the-loop simulation investigating the feasibility of the concept elements (CE) under consideration as part of the Distributed Air Ground Traffic Management (DAG-TM) program. This work was completed as part of the Advanced Air Transportation Technologies (AATT) project under NASA's Airspace Systems program. The activities compared the operational benefits and the feasibility of the CEs with respect to a baseline environment. This document is the final report that describes the experimental method, results, and lessons learned from the September 2002 activities.

1.1 BACKGROUND

1.1.1 Distributed Air Ground Traffic Management

DAG-TM is a proposed solution for expanding airspace capacity limits. It alters the roles and responsibilities of the stakeholders to permit more user-preferred routing, increased flexibility, increased system capacity, and improved operational efficiency. DAG-TM is based on the fundamental premise that all system participants can be information suppliers and users, thereby enabling collaboration and/or distribution at all levels of traffic management decision-making. This new environment will achieve successful operation through new human-centered operational paradigms enabled by procedural and technological innovations. These innovations include:

- Decision support tools (DST)
- Information sharing
- Communication, navigation, and surveillance (CNS)/air traffic management (ATM) technologies.

The present study investigated the following three DAG-TM concept elements: *En Route Free Maneuvering*, *En Route Trajectory Negotiation*, and *Terminal Arrival: Self-Spacing for Merging and In-Trail Separation*.

1.1.1.1 Concept Element 5

In CE 5, *En Route Free Maneuvering*, appropriately equipped aircraft accept the responsibility to maintain separation from other aircraft while exercising the authority to freely maneuver in en route airspace. Free maneuvering aircraft have the authority to establish new user-preferred trajectories with the restriction that new trajectories conform to local traffic flow management (TFM) constraints and do not create traffic conflicts. The flight crew's role is to avoid conflicts with other aircraft or airborne hazards (e.g., special use airspace, weather) by maintaining separation while meeting a required time of arrival (RTA). Free maneuvering aircraft have DSTs that enable situation awareness, allow flight crews to maintain separation from other aircraft without air traffic service provider (ATSP) assistance, and provide trajectory planning capabilities (Phillips, 2000).

1.1.1.2 Concept Element 6

CE 6, *En Route Trajectory Negotiation*, focuses on real-time collaboration among the DAG-TM stakeholders (flight crews, ATSP, Airline Operations Center [AOC]) for implementing enhanced user-preferred trajectory changes. A trajectory change may be initiated by any of the stakeholders, but responsibility for separation remains with the ATSP (Couluris, 2000). The flight crew, while given authority to negotiate trajectory changes, conforms (if equipped) to TFM constraints defined by the ATSP. Because these requests will increasingly conform to TFM constraints, ATSPs should be able to accommodate them more often without increasing their workload.

Implicit to the concept is the coordination between sector controllers for efficient, integrated flight planning. This requires a trajectory-oriented approach that will enable controllers to plan and coordinate trajectories across sector boundaries while maintaining separation and conforming to TFM constraints. The AOC may specify airline constraints and preferences (related to fuel efficiency, scheduling, or passenger comfort) and may initiate both long- and short-term trajectory changes. In the current concept description, the AOC-defined constraints and preferences are transmitted to the ATSP and flight deck (FD).

1.1.1.3 Concept Element 11

CE 11, *Terminal Arrival: Self-Spacing for Merging and In-Trail Separation*, is intended to increase terminal area throughput by providing flight crews and ATSPs with a method for reducing in-trail separation (Sorensen, 2000). Time-based spacing, rather than distance-based, allows for spacing compression as aircraft speeds decrease throughout the approach. Appropriately equipped aircraft entering an arrival stream are assigned an in-trail spacing interval behind a lead aircraft. The ATSP specifies the interval and lead aircraft and monitors traffic flow and maintains separation. Flight crews receive traffic intent information through airborne DSTs that support the merging and spacing operations.

2. METHOD

2.1 SIMULATION FACILITIES

The DAG-TM simulation environment was distributed among different facilities and laboratories at NASA ARC. Three main laboratories were involved:

- Airspace Operations Laboratory (AOL) providing aircraft target generation, air traffic control (ATC) and management stations augmented with a Center Terminal Radar Approach Control (TRACON) Automation System (CTAS) decision support tools, and the Multi Aircraft Control System (MACS)
- Flight Deck Display Research Laboratory providing mid-fidelity desktop simulators equipped with cockpit situation display (CSD)
- Crew Vehicle Systems Research Facility (CVSRF) providing the Advanced Concepts Flight Simulator (ACFS), a high fidelity, full mission flight simulator.

Figure 1 illustrates the DAG-TM simulation architecture. All major components of the simulation were connected via the Aeronautical Data link and Radar Simulator (ADRS) processor. The ADRS functioned as the communication management and data distribution hub. It also simulated a controller pilot data link communication (CPDLC) system by receiving data link information from simulated aircraft or ground facilities in different formats and then delaying, converting, and forwarding the information as required. See Prevot, Palmer, Smith, and Callantine (2002) for a detailed description of the DAG-TM simulation environment at NASA ARC.

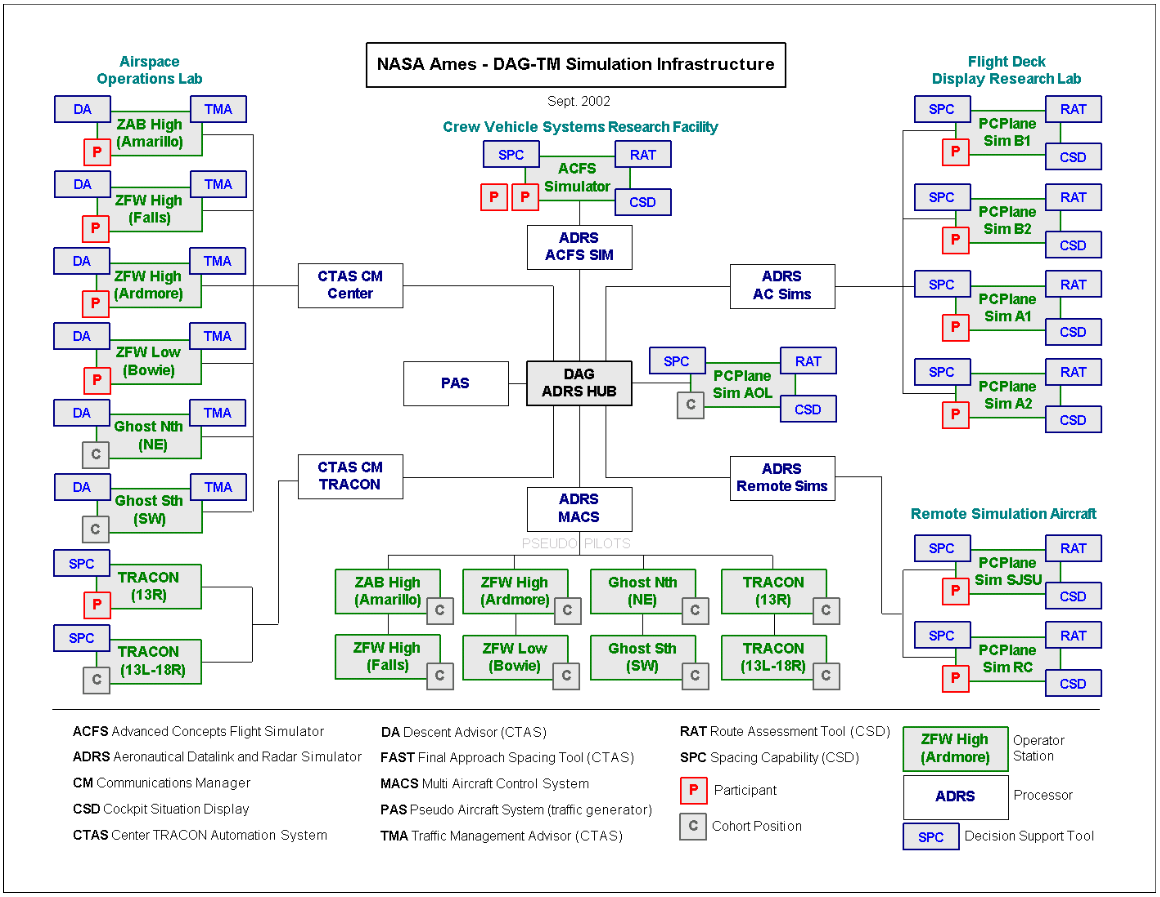

Figure 1. NASA ARC Distributed Air-Ground Simulation Architecture.

2.1.1 Airspace Operations Laboratory

The AOL is a multifunctional simulation laboratory. It controlled the overall progress of the simulation, hosted the air traffic control and management facilities, and piloted the majority of the aircraft throughout the scenario. The air traffic control facilities consisted of several controller workstations distributed through two control rooms (Figure 2). Each workstation was configurable to an en route sector, center airspace, or TRACON airspace and was equipped with air-ground/ground-ground communications lines. Workstations used modified plan-view graphical user interfaces (PGUI) that were part of the CTAS. MACS control stations were situated in a separate room within the AOL. See Section 2.1.1.1 for a description of the MACS.

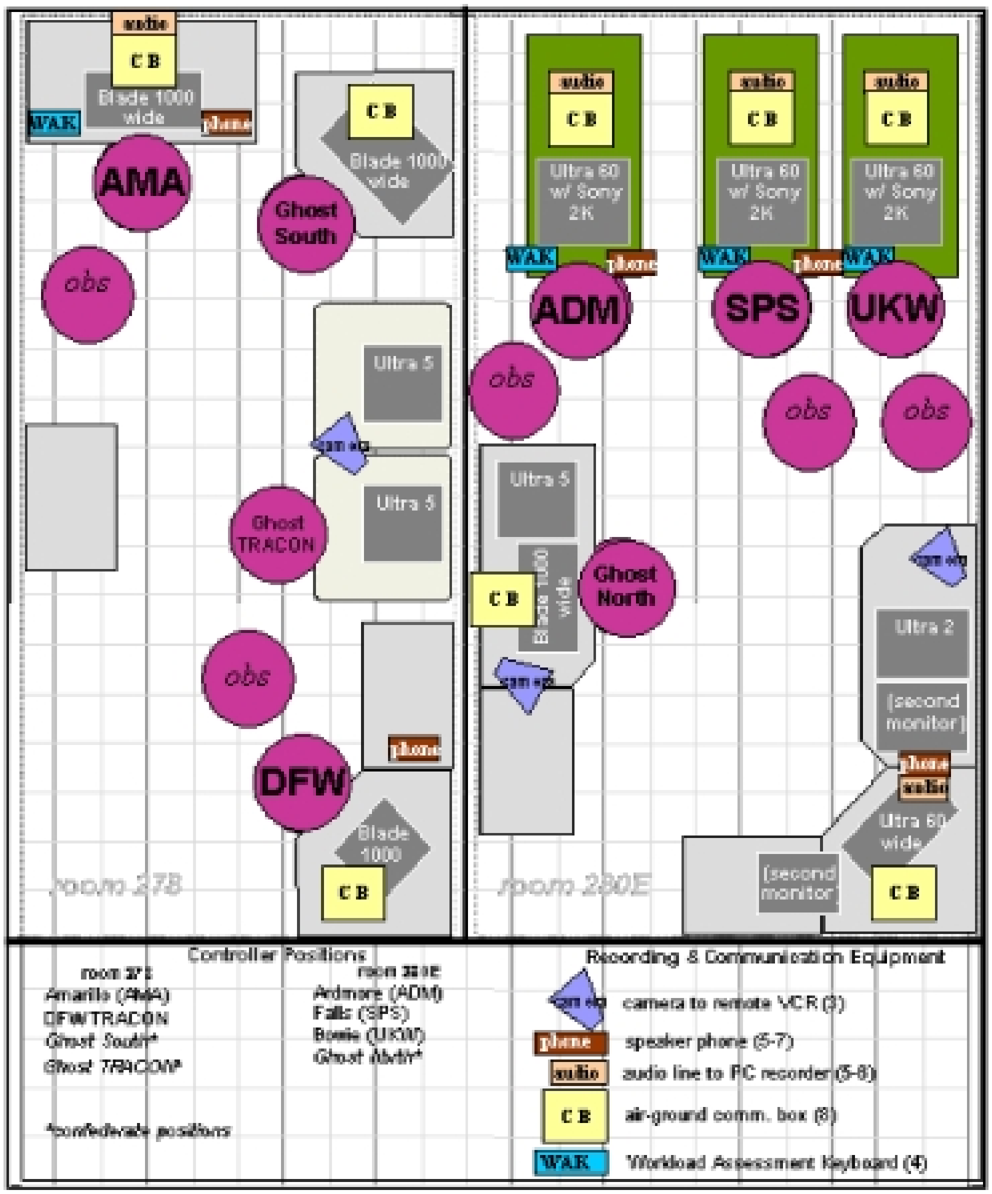

Figure 2. Airspace Operations Laboratory (AOL) Layout.

Figure 2 shows the AOL layout for the present study. Fort Worth Center (ZFW) positions (Ardmore, Wichita Falls, and Bowie) and the Ghost North positions were located in room 280E. Amarillo, Dallas/Fort Worth TRACON, and Ghost South positions were located in room 278.

The communication system included air-ground communication units at each controller position. The communication units provided two-way voice communications between controllers and pilots. The communication equipment also could have been used for ground-ground voice communications with other controllers. Conventional telephone lines provided a second means for ground-ground voice communications. Telephones were located either at individual positions or were shared between two positions.

The AOL was equipped with several types of recording equipment. Cameras, connected to a remote videocassette recorder (VCR), were used to capture the activities of subject controllers. An audio recording system was capable of capturing all voice communications transmitted via the communication boxes. Workload assessment keyboards (WAK) were used to record real-time workload ratings.

2.1.1.1 Multi Aircraft Control System

The MACS is a mid-fidelity, multi-aircraft control station (Figure 3). The display provided several flight deck components:

- Mode control panel (MCP)
- CSD
- Primary flight display (PFD)
- Flight management system (FMS) route and vertical navigation (VNAV) panels
- Self-spacing panel.

The MACS also maintained a list of all aircraft that were controlled from the station and a “to do” list that tracked those aircraft for which an operator action, like a radio check-in or a lower altitude setting, was expected. When actions were needed, MACS prompted operators by highlighting aircraft in list windows (see Figure 3).

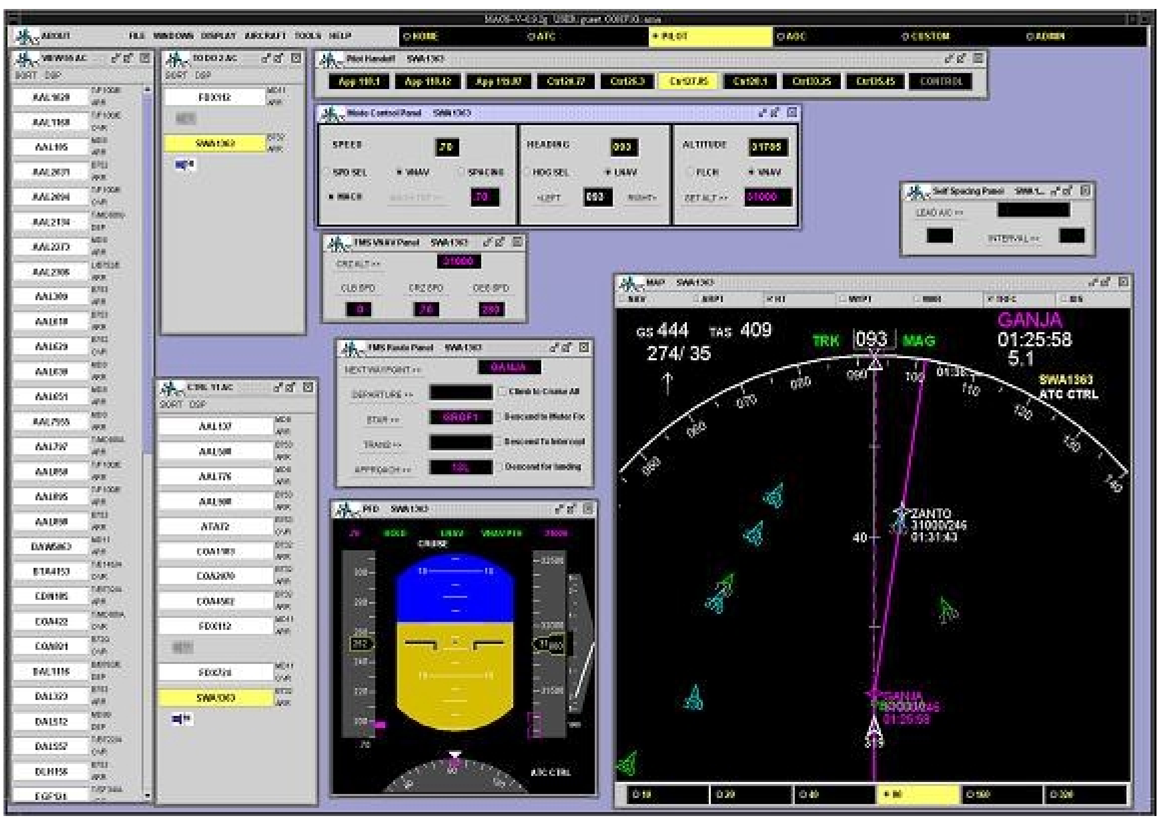

Figure 3. The Multi Aircraft Control System Interface.

The operator could enter basic autopilot commands on the MCP and could enter lateral navigation (LNAV) and VNAV commands on the FMS route panel and FMS VNAV panel, respectively. A pilot handoff panel allowed the operator to hand the aircraft to another MACS operator controlling aircraft on a different frequency.

2.1.2 Flight Deck Display Research Laboratory

The Flight Deck Display Research Laboratory housed the PCPlane desktop simulators. PCPlane is a mid-fidelity, single-aircraft flight deck simulator with a specialized CSD and a custom data link interface. The lab has the capability for multiple workstations to participate as additional flight decks in traffic scenarios. This capability increased the potential for free maneuvering aircraft participants as required for DAG-TM research.

2.1.2.1 PCPlane Flight Deck Simulator

PCPlane, developed at NASA Langley Research Center, is a personal computer (PC)-based simulator of a Boeing 757 flight deck (Figure 4). The simulator included a MCP, PFD with CSD (see Section 2.1.2.2), FMS, and data link window. It interfaced to the ADRS.

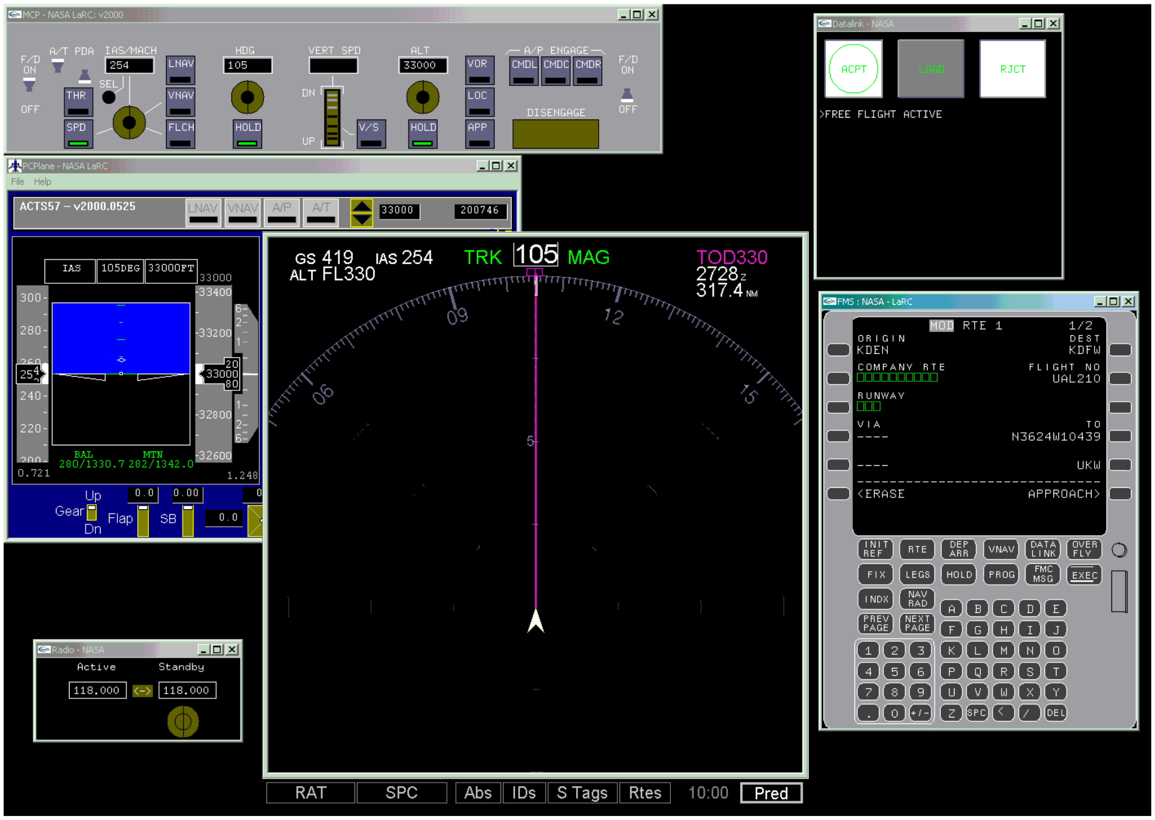

Figure 4. PCPlane Flight Deck Simulator.

2.1.2.2 Cockpit Situation Display

The CSD, developed at NASA ARC, allowed the flight crew to operate within a distributed environment by providing several specialized flight management tools:

- Primary alerting system (PAS)
- Route analysis tool (RAT)
- Terminal area approach spacing tool.

The PAS alerted the flight crew to potential traffic conflicts along the current flight path. It was designed to detect any loss of separation (LOS) between own ship and another aircraft. PAS provided three levels of conflict alerting. The RAT allowed the flight crew to develop and test potential flight plan changes before they were actually implemented (Figure 5). The approach spacing tool (Figure 6) allowed the flight crew to self-space behind a lead aircraft by selecting the lead and a time or distance to follow. The appropriate speed to achieve/maintain the desired spacing was computed and indicated through display symbology. In-trail spacing used an airborne inter-arrival spacing algorithm developed by NASA Langley Research Center to aid pilots in maintaining a time-based spacing interval (Abbott, 2002). See Raytheon Air Traffic Management System Development and Integration (ATMSDI) Team (2002) for a detailed description of the NASA Ames CSD.

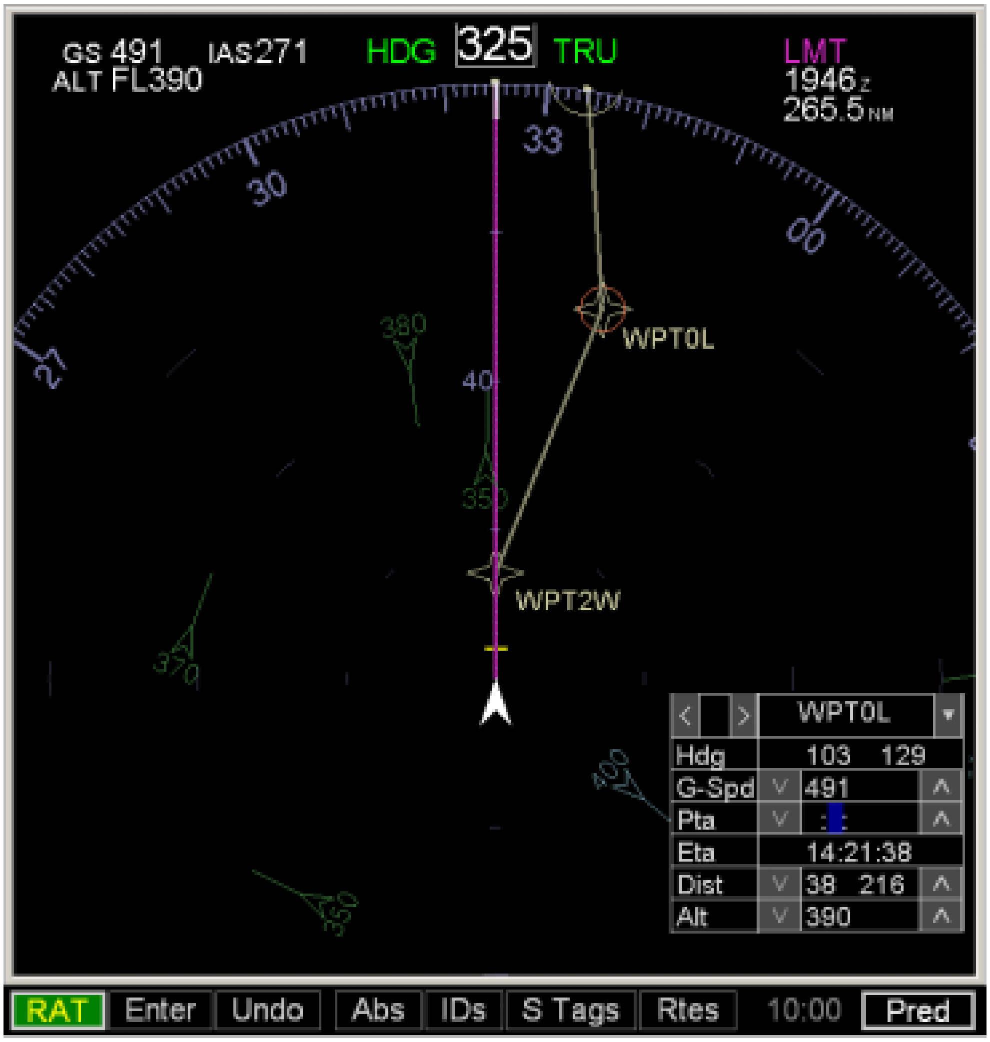

Figure 5. CSD with Route Analysis Tool Engaged.

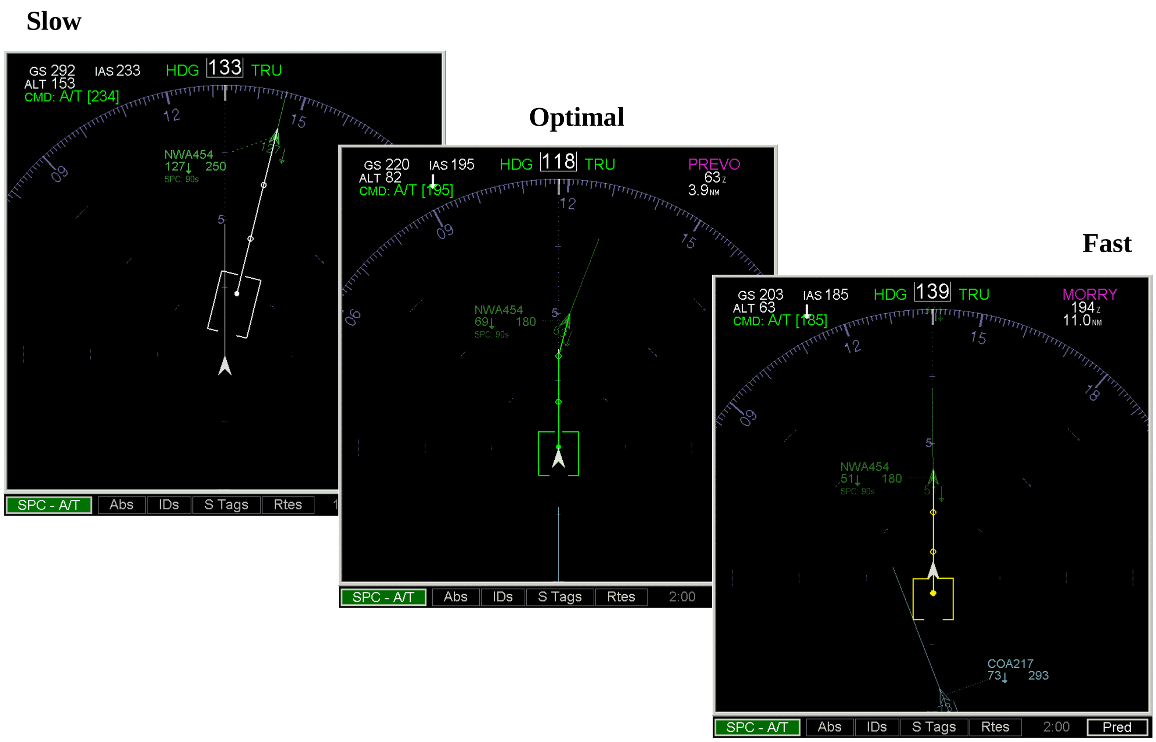

Figure 6. CSD with Approach Spacing Tool Engaged.

2.1.3 Crew Vehicle Systems Research Facility

The CVSRF houses the ACFS used in the DAG-TM simulation environment. It is a 6-degree of freedom, full-mission flight simulator equipped with future air navigation system (FANS)-type data link capabilities and CSDs on both the captain's and the first officer's position. CSD functions and capabilities were identical to those of the PCPlane workstation.

2.1.4 Remote Pilot Workstations

Two PCPlane-based pilot workstations were set up at remote locations – one at the Rotorcraft Laboratory on the grounds of ARC and the other at San Jose State University.

2.2 Participants

2.2.1 Subject Participants

Five full-performance level controllers participated in the study. All five controllers had previously participated in DAG-TM events at NASA Ames. The controllers' years of experience ranged from 12 to 24, with a mean of approximately 18. The controllers worked four en route sector positions and one TRACON position. All of the controllers were quite familiar with the sector that they controlled during the simulations (mean of 4.4 on a 5-point rating scale, 1 = very unfamiliar, 5 = very familiar).

Eight air transport rated pilots participated in the study. Only one of the eight pilots had previous experience with the DAG-TM concepts. The pilots ranged in age from 32 to 64 years, with a mean of 53 years. Total number of flight hours for each pilot ranged from 3,000 to 20,000, with a mean of approximately 11,650 hours. All eight pilots had glass cockpit experience, with a group mean of approximately 3,200 total hours with the glass cockpit. One pilot team (first and second officer) flew the ACFS and the six other pilots individually flew PCPlane desktop flight deck stations.

2.2.2 Support Participants

The support participants, those who participated in the experiment on behalf of the experimenters but were not test subjects, consisted of three confederate controllers and 10 pseudo-pilots. The three controllers (retired ATSPs) worked two adjacent en route sector positions and a TRACON position. The pseudo-pilots (general aviation licensed and/or student pilots) worked the MACS control stations. Each pseudo-pilot, or pair of pseudo-pilots, flew all aircraft within a designated airspace (which corresponded to controller airspace sectors).

2.3 OPERATIONAL ENVIRONMENT

2.3.1 Airspace

The simulation traffic scenarios occupied the airspace that handles northwest arrivals into Dallas/Fort Worth International Airport (DFW). The airspace encompassed several en route sectors in ZFW and a portion of the Albuquerque Center (ZAB) airspace (Figure 7). DFW and Dallas Love Field (DAL) arrivals transitioned to the DFW TRACON. Transition occurred at the TRACON boundary fix, BAMBE for DFW arrivals and GREGS for DAL arrivals. The UKW merge point fed both boundary fixes. BAMBE was the meter fix for all DFW arrivals in this simulation.

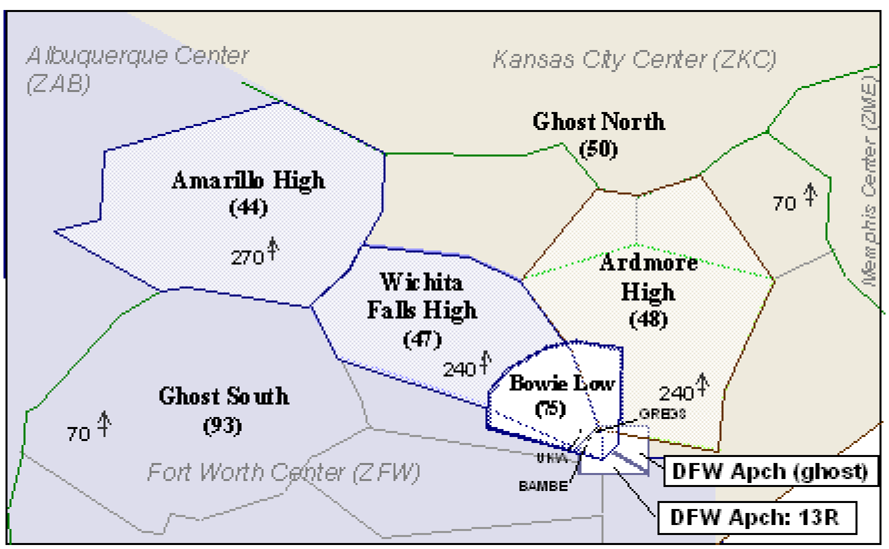

Figure 7. Simulation Airspace.

The five airspace sectors staffed by subject controllers:

- Amarillo — a high-altitude sector within ZAB airspace
- Wichita Falls — a high-altitude sector within ZFW airspace that handled a large portion of arrivals to DFW
- Ardmore* — a high-altitude sector within ZFW airspace that typically handled a mix of DFW arrivals/departures and over-flights

- Bowie — a low-altitude sector within ZFW airspace that handled northwest arrivals to DFW. The merge point (UKW) and northwest TRACON boundary fix (BAMBE) was within Bowie airspace.
- TRACON — Terminal Radar Approach Control for DFW runway (RW) 13R.

* The Ardmore sector's northern boundary was modified for this study, extending north to the ZFW/Kansas City Center (ZKC) boundaries. This was done to enable arrival aircraft from the north to enter one of the test sectors earlier in the scenario.

Confederate controllers staffed three other positions:

- Ghost North — a "super" en route (high/low altitude) sector that included the northern part of ZFW and a southern portion of ZKC
- Ghost South — a "super" en route (high/low altitude) sector that included the southwestern part of ZFW and an eastern portion of ZAB
- Ghost TRACON — a Terminal Radar Approach Control position that handled traffic for DFW RW18R, DFW 13R Tower, and DAL RW 13L.

2.3.2 Air Traffic Service Provider Tools

En route and TRACON controllers were provided with a variety of CTAS tools, such as the Traffic Management Advisor (TMA) and the Descent Advisor (DA). TMA optimizes the arrival flow by scheduling aircraft on a first-come, first-served basis. Aircraft with earlier estimated times of arrival (ETA) are assigned earlier scheduled times of arrival (STA). The TMA scheduler was set up, for this study, to provide a minimum of 7 miles-in-trail spacing at the meter fix. At BAMBE, with an 11,000 foot and 250 knot crossing restriction, this translated to roughly 82 seconds minimum spacing between aircraft. The DA provided speed advisories, supported manual route planning, and calculated RTAs.

Controller workstations included:

- Controller PGUI
- CTAS DSTs
  - TMA with ETA and STA capability
  - Conflict detection
  - Trial (route) planning
  - DA with RTA capability
- Data link capability
- Control exchange (autonomous to managed and vice versa) capability
- Arrival spacing capability (time-based)

To support DAG-TM CE 11, airborne self-spacing functionality was available to the TRACON controller through the display of actual and advised spacing intervals. History circles supported the airborne self-spacing function by indicating the desired position of a trailing aircraft behind a lead aircraft.

2.3.3 Summary of Operational Parameters

Several parameters defined the operational environment. These parameters (e.g., an aircraft's equipage and the level of distributed control) determined the extent to which each aircraft operated within a specified airspace. Table 1 summarizes the operational parameters. Each parameter is further explained in the following four sections: Aircraft Equipage (Section 2.3.4), Levels of Distributed Control (Section 2.3.5), DAG-TM Capabilities (Section 2.3.6), and Modes of Operation (Section 2.3.7).

Table 1. Summary of Operational Parameters.

Modes of Operation	Aircraft Equipage	En Route Airspace		Terminal Airspace	
		Level of Control	DAG-TM Capabilities	Level of Control	DAG-TM Capabilities
Baseline	Equipped	Managed	None	Managed	None
	Unequipped	Managed	None	Managed	None
CE 6 & 11	Equipped	Managed w/ negotiation authority	Trajectory negotiating	Managed w/ self-spacing authority	In-trail spacing
	Unequipped	Managed	None	Managed w/ self-spacing authority	In-trail spacing
CE 5 & 11	Equipped	Autonomous	Free maneuvering	Managed w/ self-spacing authority	In-trail spacing
	Unequipped	Managed	None	Managed w/ self-spacing authority	In-trail spacing

2.3.4 Aircraft Equipage

Pilots participating in the study were given tools that allowed them to operate in a DAG-TM-type environment. A mix of equipped and unequipped aircraft flew in the airspace. Approximately 90 unequipped and 7 equipped aircraft participated in each 1-hour simulation run.

2.3.4.1 Equipped Aircraft

Equipped aircraft were provided with technology necessary to create and execute user-preferred trajectory changes within en route airspace, detect and resolve potential traffic conflicts, and maintains self-spacing within terminal airspace. Equipped aircraft possessed the following:

- FMS
- CSD with
  - Aircraft intent information
  - Conflict detection with manual resolution
  - Route planning
  - RTA capability
  - Precision descent procedure (a procedure flown with conventional FMS VNAV)
  - Arrival spacing capability (time-based)
- Automatic dependent surveillance-broadcast (ADS-B)
- CPDLC

2.3.4.2 Unequipped Aircraft

Unequipped aircraft simulated current-day capabilities with a few additional ones. Unequipped aircraft possessed the following:

- FMS
- CSD with
  - Traffic Alert and Collision Avoidance System (TCAS)
  - RTA capability
  - Arrival spacing capability (time-based)
- ADS-B
- CPDLC

2.3.5 Levels of Distributed Control

The level of distributed control refers to the delegation of control between the ATSP and the flight crew, namely who maintains separation responsibility? It also defines the extent to which the flight crew can self-manage their flight. For example, the flight crew of an autonomous aircraft was permitted to make flight plan changes without ATSP approval.

2.3.5.1 Managed

An aircraft was managed if the controller retained separation responsibility for that aircraft. Unequipped aircraft were managed at all times. An equipped aircraft was managed if the flight crew had not yet been given autonomous control, or if the flight crew of an autonomous aircraft requested (and the controller accepted) that ATC resume separation responsibility. In the case of a potential conflict between a managed and autonomous aircraft, the autonomous aircraft and the controller acting on behalf of the managed aircraft followed specified flight rules.

2.3.5.2 Managed With Negotiation Authority

Managed aircraft with negotiation authority followed the same rules as managed with one exception. The flight crews of equipped aircraft were given authority to negotiate trajectory changes (see Section 2.3.6.1). All trajectory changes had to be submitted to the ATSP for approval before the change could be implemented, and negotiation could take place only while in en route airspace.

2.3.5.3 Managed With Self-Spacing Authority

Managed aircraft with self-spacing authority followed the same rules as managed with one exception. Upon entering terminal airspace, the flight crews of both equipped and unequipped aircraft were given authority to implement in-trail spacing along an FMS arrival route (see Section 2.3.6.3). Each aircraft's flight crew was responsible for maintaining the specified temporal spacing. The responsibility of maintaining the separation minimum was always that of the ATSP. Self-spacing took place only while within terminal airspace.

2.3.5.4 Autonomous

An aircraft was autonomous and allowed to free maneuver (Section 2.3.6.2) if it was appropriately equipped and had been granted the authority to execute user-preferred trajectory changes without requesting ATC clearance. The flight crew of an autonomous aircraft had the responsibility of maintaining the separation minimum and resolving potential traffic conflicts while following the rules-of-the-road (see Section 2.3.8). Autonomous control was granted only while in en route airspace.

2.3.6 DAG-TM Capabilities

2.3.6.1 Trajectory Negotiating

The flight crew generated a user-preferred trajectory modification that was data-linked to the ATSP. The ATSP reviewed the request, either accepting or rejecting it, and data-linked the response back to the flight crew. If accepted, the flight crew then instructed the aircraft's FMS to initiate the trajectory. On-board automation broadcasted the modified trajectory using ADS-B to the ATSP and to other aircraft. It was expected that trajectory modifications would not create near-term conflicts with other aircraft in the vicinity.

The ATSP provided the flight crews with local TFM constraints such as RTA, altitude, speed, or spacing restrictions, and could up-link ATSP-generated clearances.

2.3.6.2 Free Maneuvering

The flight crew generated user-preferred trajectory modifications and instructed the aircraft's FMS to initiate the trajectory. On-board automation broadcasted the modified trajectory using ADS-B to the ATSP and to other aircraft. The flight crew was responsible for ensuring that trajectory changes did not generate near-term conflicts with other aircraft in the vicinity. A conflict detection tool provided predicted conflict alerts that required the flight crew to respond accordingly, either taking evasive action or allowing the intruder aircraft to maneuver. The manner in which a traffic conflict was resolved was determined by the rules-of-the-road.

2.3.6.3 In-Trail Spacing

Upon entering terminal airspace, the flight crew flew along a structured arrival route behind a designated lead aircraft. The ATSP was responsible for assigning the lead aircraft and the temporal spacing interval. The flight crew was responsible for maintaining the assigned spacing interval. DSTs allowed the flight crew to monitor their spacing performance and compensate for spacing variations within the arrival stream. The lead aircraft was always in the same arrival stream.

2.3.7 Modes of Operation

This study investigated three modes of operation – a baseline mode and two DAG-TM modes. The Baseline mode was designed to equate to current-day capabilities and tools. In all modes of operation, controllers were attempting to deliver aircraft to the BAMBE meter fix at the TMA scheduled times.

2.3.7.1 Baseline

The Baseline condition was designed to represent present-day airspace operations with the addition of the TMA. Using current standard operating procedures (SOP), aircraft traveled through en route airspace towards an arrival airport. All aircraft had ADS-B capability, which permitted the display of proximal traffic on the CSD. In this mode of operation, the pilots could use the CSD only for situation awareness. They transitioned to TRACON airspace and landed with the assistance of ATSPs. All aircraft, whether equipped or unequipped, were managed; and ATSPs were responsible for maintaining separation.

Technologies

Controller Workstations

- CTAS TMA (only sequence list and delay information)

Unequipped Aircraft

- As defined in Section 2.3.4.2, except…
- No arrival spacing capability
- No CPDLC
- No RTA capability

Equipped Aircraft

- All equipped aircraft function as unequipped aircraft

Standard Operating Procedures

- Current-day SOPs
- ATSP are expected to conform to RTAs
- All aircraft fly FMS Transition in TRACON

2.3.7.2 Trajectory Negotiation (CE 6 & CE 11)

Aircraft flew en route towards an arrival airport. All aircraft were under the control of the ATSP and had data-linked meter fix times. A few aircraft were provided with the tools that enabled them to modify flight path trajectories during the en route phase. These aircraft could send and receive trajectory requests to and from the ATSP. Delays resulted in aircraft being taken off their preferred routes. Less-than-optimal routing motivated the flight crews of equipped aircraft to request alternative trajectories.

One or more streams of traffic merged at the TRACON meter fix and entered terminal airspace. ATSP issued a limited delegation clearance to both equipped and unequipped aircraft. The clearance included a time-based spacing interval and lead aircraft to space behind. As many DFW arrivals as possible landed on RW13R, and Runway 18R was available for excess aircraft. A few aircraft were routed to DAL RW13L. Using flight deck DSTs, the flight crews identified the lead aircraft and then used the assigned spacing interval to self-space behind it.

The TRACON controllers monitored the air traffic situation and used a ground-based DST to monitor conformance and assign spacing intervals. The controller managed aircraft by providing sequencing and spacing guidance while providing separation assurance for all aircraft.

Technologies

Controller Workstations

- As defined in Section 2.3.2, except…
- No control exchange capability

Unequipped Aircraft

- As defined in Section 2.3.4.2, except…

Equipped Aircraft

- As defined in Section 2.3.4.1

Standard Operating Procedures

- ATSP maintains separation responsibility for all aircraft at all times
- ATSP attempts to maintain lateral trajectories by first considering speed and altitude changes
- Only equipped aircraft may request trajectory changes
- All trajectory change requests must be conflict-free
- ATSP has the authority to reject any trajectory change request
- RTA and speed advisories are automatically up-linked to arriving equipped aircraft
- Equipped aircraft are responsible for meeting their RTAs (± 15 seconds) and are expected to stay on their 4D trajectories
- Modifications to speed (by equipped aircraft) to maintain RTA do not need concurrence
- Speed clearances from the ATSP supercede the current up-linked RTA
- All aircraft follow the precision descent procedure
- ATSP may assign spacing interval and lead aircraft upon entry into terminal airspace
- Flight crew has the authority to reject the spacing clearance
- All aircraft fly FMS transition in TRACON

2.3.7.3 Free Maneuvering (CE 5 & CE 11)

Using flight deck DSTs, flight crews of equipped aircraft had the authority to execute flight plan changes without ATSP review. Autonomous aircraft were responsible for maintaining separation from traffic. Autonomous aircraft were required to use cockpit DSTs for conflict detection and resolution while meeting RTAs and conforming to the rules-of-the-road (see Section 2.3.8.1). Limited exchange of control between equipped aircraft and ATSP occurred. All aircraft transitioned to managed TRACON airspace.

The en route controllers monitored the air traffic situation and used ground-based DSTs to compute trajectories and broadcast RTAs for terminal airspace transition. Controllers managed unequipped aircraft (and equipped aircraft if requested) and provided separation from autonomous and other managed aircraft as well as airspace hazards. The rules-of-the-road defined the responsibilities of controller and flight crew for resolving traffic conflicts. The low-altitude en route controller was responsible for transitioning free maneuvering aircraft from autonomous to managed control. Entry of autonomous aircraft in terminal airspace adhered to the procedures for transitioning (see Section 2.3.8.3).

As with *DAG-TM Trajectory Negotiation* mode, one or more streams of traffic merged at the TRACON meter fix and entered terminal airspace. The TRACON controller managed the arrival stream through sequencing and spacing guidance (i.e., spacing interval and lead aircraft) while providing separation assurance for all aircraft.

Technologies

Controller Workstations

- As defined in Section 2.3.2

Unequipped Aircraft

- As defined in Section 2.3.4.2.

Equipped Aircraft

- As defined in Section 2.3.4.1

Standard Operating Procedures

- ATSP maintains separation responsibility for all managed aircraft
- ATSP attempts to maintain lateral trajectories (for managed aircraft) by first considering speed and altitude changes
- Autonomous aircraft are responsible for maintaining separation from other aircraft and airspace hazards
- ATSP may not issue clearances/instructions to autonomous aircraft except for terminal airspace entry
- Autonomous aircraft are responsible for meeting their RTAs (± 15 seconds) and are expected to stay on their 4D trajectories
- ATSP may only cancel autonomous operations by exception (e.g., if an autonomous aircraft is off its RTA)
- RTA and speed advisories are automatically up-linked to arriving autonomous aircraft
- Traffic conflicts are resolved by following the rules-of-the-road
- Autonomous aircraft are free to modify 4D trajectories without consulting the ATSP
- Autonomous aircraft may request cancellation of autonomous operations at any time (if complying to standard ATC handoff rules)
- Cancellation of autonomous operations must be approved by the ATSP
- Equipped aircraft with autonomous control canceled remain under control of the ATSP until ATSP clears the aircraft to resume control (either by request or concurrence)
- Autonomous aircraft may request ATC assistance (for conflict resolution, flow control, and traffic management considerations)
- ATSP is expected to assist the flight crew of autonomous aircraft with request (if request is received in a timely manner)
- ATSP may point out potential traffic conflicts to autonomous aircraft (time and workload permitting)
- ATSP may cancel autonomous control if a conflict with a managed aircraft cannot be resolved otherwise
- All aircraft fly FMS transition in TRACON
- ATSP assigns spacing interval and lead aircraft upon entry into terminal airspace
- Equipped aircraft have the authority to reject spacing assignment

2.3.7.4 Summary of Modes of Operations

Table 2 and Table 3 summarizes the equipment and tools that were available to the flight crews and ATSPs for each mode of operation.

Table 2. Aircraft Equipment for Modes of Operation.

Aircraft Equipment	Baseline	CE 6 & 11 Trajectory		CE 5 & 11 Free Maneuvering	
	All Aircraft	*Unequipped*	*Equipped*	*Unequipped*	*Equipped*
FMS	x	x	x	x	x
RTA		x	x	x	x
CSD	x	x	x	x	x
ADS-B	x	x	x	x	x
CPDLC		x	x	x	x
Conflict detection			x		x
Route planning			x		x
Precision descent procedure		x	x	x	x
Arrival spacing capability		x	x	x	x
Rules-of-the-road					x

Table 3. ATSP Tools for Modes of Operation.

ATSP Tools	Baseline	CE 6 & 11 Trajectory Negotiation	CE 5 & 11 Free Maneuvering
TMA	x	x	x
Descent advisor		x	x
Conflict detection		x	x
Trial planning		x	x
Arrival spacing capability		x	x
Control exchange capability		x	x
Rules-of-the-road			x

2.3.8 Rules-Of-The-Road

A set of flight rules was developed to define the responsibilities and actions of aircraft involved in traffic conflicts. These rules applied to situations involving autonomous aircraft and, thus, were in effect only during the *Free Maneuvering (CE 5 & CE 11)* mode of operation. The rules-of-the-road were specific to the different types of airspace, as described below.

2.3.8.1 En Route Airspace

Aircraft were normally separated during en route operations by flying a correct altitude for direction of flight.

Flight levels (FL) below 290:

- On headings of 360° to 179° (i.e., eastbound), aircraft must fly at odd numbered altitudes or FLs (e.g., 170, 190, 210).
- On headings of 180° to 359° (i.e., westbound), aircraft must fly at even numbered altitudes or FLs (e.g., 180, 200, 220).

FLs 290 and above:

- On headings of 360° to 179° (i.e., eastbound), aircraft must fly the following odd numbered altitudes or FLs – 290, 330, 370, etc.
- On headings of 180° to 359° (i.e., westbound), aircraft must fly the following odd numbered altitudes or FLs – 310, 350, 390, etc.

The general rule for encounters between two aircraft (at least one being under autonomous control) is that the right-of-way was determined by flight priority:

- All arriving aircraft that are within 200 nautical miles (nm) of their arrival airport will have the right-of-way over all other aircraft, except emergency. (The arrival status will be indicated on CSD and ATC displays.)

If the encounter was between two aircraft with equal priority (e.g., between two arriving aircraft or two en route aircraft) and both were in cruise, the right-of-way was based on the following rules:

- The aircraft not flying the correct altitude for direction of flight will maneuver.

- For encounters between aircraft that are flying the correct altitude for direction of flight, the aircraft on the right will have the right-of-way.

In situations when aircraft of equal priority were climbing, descending, or overtaking, the following rules applied:

- For encounters between climbing and descending aircraft and an aircraft in level flight, the aircraft in level flight will have the right-of-way.
- For encounters between a climbing and a descending aircraft, the descending aircraft will have the right-of-way.
- For two aircraft climbing or descending, the aircraft on the right will have the right-of-way.
- For encounters where one aircraft is overtaking another on the same path (approximately 20° or less), the lead aircraft will have the right-of-way.

In addition, any aircraft could agree to cede the right-of-way to another aircraft at any time. The en route airspace rules-of-the-road applied only to encounters in which autonomous aircraft were involved. Resolution of conflicts between managed aircraft was the discretion of the ATSP. For encounters between managed and autonomous aircraft, the ATSP applied these rules on behalf of the managed aircraft.

2.3.8.2 Terminal Airspace

All aircraft were under ATSP control. No free maneuvering or trajectory change requests were allowed.

2.3.8.3 Transitioning Between En Route and Terminal Airspace

All aircraft were required to meet a RTA for the initial transition from en route to terminal airspace. If an autonomous aircraft failed to meet its RTA, the low-altitude en route controller (Bowie) could cancel autonomous operations and control the aircraft so as to provide an acceptable flow to the TRACON.

2.3.9 Scenario Description

Four traffic scenarios were developed for the simulation. Each 60-minute scenario depicted a DFW arrival push. Traffic density was medium to heavy, with approximately 90 aircraft per scenario. The scenarios, having similar characteristics and parameters, were considered equivalent in traffic flow and complexity. One scenario was used during the training runs and the other three during data collection runs. The three data collection scenarios were designated DAG 1, DAG 2, and DAG 3.

Scenario events included handling conflicts, routing DAL arrivals through DFW arrivals, merging aircraft into an arrival stream in Bowie, and coping with traffic-load delays in the arrival flow. Approximately 40 aircraft were involved in conflicts that had been built into the scenario. Winds were consistent across all scenarios.

The three data collection scenarios are described in the following paragraph and in Table 4. The scenarios were set in the northwestern section of ZFW airspace and depicted a single-gate arrival flow of approximately 7 minutes in-trail.

Table 4. Summary of Traffic Scenarios Used for Data Collection.

Parameter	Traffic Scenario		
	DAG 1	DAG 2	DAG 3
Total number of aircraft	88	94	87
Number of arrivals	46 (42 + 1 NASA UKW7, 4 DAL)	50 (47 UKW7, 3 DAL)	45 (42 UKW7, 2 DAL)
Number of departures	14	13 (12 DFW, 1 DAL)	13 (12 DFW, 1 DAL)
Number of over-flights	28	31	30
Arrival traffic flow	From northwest, north, and northeast to DFW	Same	Same
Arrival	UKW7 (via HEATR, GANJA, SPS, MOOSE) GREGS4	Same	Same
Departure traffic flow	Mostly to the north through Ardmore	Same	Same
Departure	TEX7	Same	Same
Over-flight traffic flow	- Start east of DFW center going to west - Start southwest of DFW center going to northeast - Start southeast of center going to northwest	Same	Same
Sector load	44 Amarillo, 47 Wichita Falls, 75 Bowie, 48 Ardmore, 248 TRACON, 50 Ghost North, and 93 Ghost South	Same	Same
Insertion point of aircraft	Mostly in Ghost sectors for those aircraft that get initialized late in scenario	Same	Same
Frequency of delays	High at end of scenario	Low	Moderate
Conflicts	Approximately 30-40 aircraft involved	Same	Same
Aircraft type	Standard jets 737, F100, 747, 757, A320, etc.	Same	Same
Suggested order of running	Third	First	Second

The majority of aircraft arrived into DFW, flying the UKW7 STAR, through the northwest cornerpost and landed on runways 13R and 18R. In addition, several northbound departures exited the scenario through Ardmore and Ghost North, and over-flight from the south, east, and west. There were four aircraft on the GREGS4 STAR to DAL RW13L. These flights had to cross the DFW UKW arrival flow.

2.4 Experimental Design

2.4.1 Independent Variable

The only independent variable in the study was the mode of operation (Table 5), as described in Section 2.3.7. The study, therefore, used a single-factor design, which varied along three levels. The conditions were designated as shown in Table 5:

Table 5. Independent Variable.

Condition	Mode of Operation
Baseline	Baseline (Section 2.3.7.1)
CE 6	Trajectory Negotiating and In-Trail Spacing (Section 2.3.7.2)
CE 5	Free Maneuvering and In-Trail Spacing (Section 2.3.7.3)

As shown in Table 6, three runs were conducted daily. Each day, participants completed a run in one of the three scenarios for each of the three operational conditions (days 1, 2, and 3 were scheduled training days). By the end of the study, participants had completed at least one run in each condition with each traffic scenario. The presentation order of the three modes of operation and the three scenarios was counterbalanced to reduce the likelihood of a learning bias.

Table 6. Presentation of Conditions and Scenarios for Data Collection Runs.

Day 4	Day 5	Day 6	Day 7
Condition-Scenario	**Condition-Scenario**	**Condition-Scenario**	**Condition-Scenario**
CE 6 - DAG 3	Baseline - DAG 2	CE 6 - DAG 1	CE 5 - DAG 3
Baseline - DAG 2	CE 5 - DAG 1	Baseline - DAG 3	CE 6 - DAG 2
CE 5 - DAG 1	CE 6 - DAG 3	CE 5 - DAG 2	Baseline - DAG 1

2.4.2 Dependent Variables

Several measures were collected during the experimental runs. Dependent variables included human performance metrics and system-level metrics. Subjective, or self-reported data, was gathered with a series of questionnaires administered after simulation runs. More precisely, participants completed flight crew or controller post-run questionnaires after every experimental run. At the end of the study, they completed flight crew or controller post-simulation questionnaires. Objective measures were collected via the simulation software recordings in both the AOL and Flight Deck Display Research Laboratory. The following two sections describe the human performance metrics and detail the system performance metrics.

2.4.2.1 Human Performance Metrics

Demographic Information

A questionnaire was used to gather demographic information and the level and type of experience of participants (controllers and pilots).

Workload

Subjective workload assessments were collected from controllers using the Air Traffic Workload Input Technique (ATWIT) (Stein, 1985). Controllers were required to rate their workload via WAKs on a scale of 1 to 7, at 4-minute intervals throughout each simulation run. Questions from the NASA-Task Load Index (TLX) were included in the controller and pilot post-run questionnaires. Other questions in the post-run questionnaires asked participants if they felt that they had sufficient time to allocate to communication and coordination tasks. In addition to self-reports, researchers made over-the-shoulder observations of controller and pilot actions. Measures of task workload, mentioned in the performance metrics section, were also collected.

Communications

During the experimental runs, the simulation software recorded ground-to-air (and air-to-ground), air-to-air, and ground-to-ground communications.

Usability, Suitability, and Acceptability

Post-run and post-simulation questionnaires addressed usability, suitability, and acceptability issues with the advanced technologies. Participants indicated whether they found their tools to be effective and usable. Participants also answered questions regarding the adequacy of the procedures and rules-of-the-road used during the simulation runs and was encouraged to make suggestions for improvement. Post-run data analysis software also calculated the number of entries made by users.

2.4.2.2 System Performance Metrics

The simulation software recorded and computed several metrics to measure system capacity, complexity, efficiency, and safety. The system level metrics is listed in Table 7 along with the relevant stakeholder. Note some metrics were not used in the analyses or reported in the result section.

Table 7. System Performance Metrics.

Metric	System Level				Stakeholder		
	Capacity	Complexity	Efficiency	Safety	En Route Controller	TRACON Controller	Flight Crew
Average distance traveled by aircraft (sector/overall)			x		x	x	x
Average time aircraft spent under control of each sector controller	x		x		x	x	
Average deviation from assigned spacing interval			x				x
Average actual time of arrival (ATA) spacing at meter fix	x		x		x		
ATA spacing at final approach fix (FAF)	x		x			x	
Average altitude deviation from crossing restriction at the meter fix		x	x	x	x		x
Number of FMS flight plan amendments		x	x		x		x
Number of altitude changes			x		x	x	x
Number of heading changes			x		x	x	x
Number of speed changes			x		x	x	x
Arrival prediction error: ∑(ATA – STA)	x		x		x	x	
Number of aircraft holds		x	x		x	x	x
Number of controller cancellations of autonomous control		x		x	x		x
Number of pilot cancellations of autonomous control		x		x	x		x
Number of conflict alert warnings				x	x	x	x
Closest point of approach (CPA) for aircraft pairs that were in conflict alert				x	x	x	x
Number of operational errors*****/LOS				x	x	x	x
Number, proportion, and nature of pilot-controller voice communications							
Number, proportion, and nature of data-linked clearances							
Number of aircraft in sector every 5 minutes	x	x			x	x	
Number of altitudes used	x	x			x	x	
Number of aircraft at each altitude	x	x			x		
Average number of entries for flight plan modification (heading, speed, and altitude)			x		x		x

2.5 Participant Training

The purpose of training was to ensure that all participants were familiar with the tools, display, procedures and rules-of-the-road that were used during the experiment. Prior to participating in data collection runs, pilot and controller subject participants trained on the DAG-TM concepts and the tools they used in the simulation. Training occurred in three phases and took place over a period of 3 days. The training schedule is outlined in Table 8.

Phase 1

The first phase included all subject participants. As a group, the pilots and controllers attended a DAG-TM orientation consisting of a description of the experiment objectives, terminology unique to the study, and detailed descriptions of the concept elements to be studied. Specific differences between CE 5 and CE 6 experimental conditions were discussed thoroughly.

Subject participants were given a description of the basic scenarios used in the simulation, including information on the airspace and traffic flows, arrival and approach charts, and descent procedure information. The pilots and controllers were provided with a brief description of the tools and the rules-of-the-road.

Phase 2

After the initial orientation, pilots and controllers were divided into separate training groups. This second phase of the training focused on user-specific instruction on the tools and procedures unique to each group.

The controllers participating in the study had previously participated in DAG-TM simulations and therefore were familiar with the concepts, terminology, and procedures. However, any changes in the displays, tools, and procedures since the previous simulations were discussed.

The pilots, however, were all new to the DAG-TM experience. To ensure that the pilots were at the same level of proficiency as controllers, researchers prepared a training packet focusing on both the concepts and tools required for participation in distributed air-ground traffic management. Researchers provided the pilots with details regarding the displays, controls, and procedures for each CE and Baseline condition. The training packet included screen shots of the CSD and diagrams of the relevant alerting symbology. Screen shots showed the steps required for specific procedures and displayed the results of selected control interactions. A section specific to PCPlane use, including communications, data link, and FMS usage, was also included.

Phase 3

The final phase of training consisted of a distributed simulation where each subject and support participant was at his or her station, and all participants completed several training runs as a group. The scenario for the training runs was designed specifically for training and was not used during any of the data collection runs. However, the training runs were still conducted in the Dallas-Ft. Worth airspace with traffic flow and airspace constraints similar to those in the data runs. Trained observers coached the participants as necessary during the training runs.

2.6 Training and Data Collection Schedule

Training and data collection occurred over a 2-week period. Table 8 and Table 9 shows the schedule for the September 2002 DAG-TM research study. The tables include the training schedule, 12 data collection runs, time slots for questionnaires and debriefings, time slot for a repeat data collection run, and travel time for participants.

Table 8. Training and Data Collection Schedule – Week 1.

	Monday 9/9/2002	**Tuesday** 9/10/2002	**Wednesday** 9/11/2002	**Thursday** 9/12/2002	**Friday** 9/13/2002
8:00	Travel Day for Participants	Introduction to DAG-TM	Briefing	Briefing	Briefing
8:30			CE 5 Training Run	Baseline Training Run	Check-in
8:45-10:00		Break			Data Collection Run 1
10:00		Experiment Briefing (CE 6, CE 5, CE 11, airspace, rules-of-the-road)	Break	Break	Break
10:15			Baseline Training Run	CE 6 Training Run	Check-in
10:30-11:45					Data Collection Run 2
12:00-1:00		Lunch	Lunch	Lunch	Lunch
1:00		Introductory Simulator Training	Debriefing	CE 5 Training Run	Check-in
1:15-2:30		Break	Break	Break	Data Collection Run 3
2:30		CE 6 Training Run	CE 6 Training Run	Debriefing	Break
3:00-4:00					Debriefing

Table 9. Training and Data Collection Schedule – Week 2.

	Monday 9/16/2002	**Tuesday** 9/17/2002	**Wednesday** 9/18/2002	**Thursday** 9/19/2002	**Friday** 9/20/2002
8:00	Briefing	Briefing	Briefing	Briefing	Travel Day for Participants
8:30	Check-in	Check-in	Check-in	Check-in	
8:45-10:00	Data Collection Run 4	Data Collection Run 7	Data Collection Run 10	Repeated Data Collection Run 5	
10:00	Break	Break	Break	Break	
10:15	Check-in	Check-in	Check-in	Post-experiment questionnaire	
10:30-11:45	Data Collection Run 5	Data Collection Run 8	Data Collection Run 11		
12:00-1:00	Lunch	Lunch	Lunch	Lunch	
1:00	Check-in	Check-in	Check-in	Check-in	
1:15-2:30	Data Collection Run 6	Data Collection Run 9	Data Collection Run 12	Demonstration Run	
2:30	Break	Break	Break	Break	
3:00-4:00	Debriefing	Debriefing	Debriefing	Debriefing	

3. RESULTS

Inferential tests were performed to compare the results from operational data across experimental conditions. However, because of the small number of subjects and restrictions typical of large-scale simulation studies, assumptions such as random sampling, independence of observations, and homogeneity of variance do not always hold. To address these difficulties, randomization tests (Byrne, 1993; Edgington, 1980; Noreen, 1989) were used to determine the significance of the inferential tests. All randomization tests, also called permutation tests, were performed with the "Resample" software (Howell, 2001). Appendix B provides a detailed description of the permutation test.

Because the performance of each flight may not have been independent from the performance of the other flights in the same run the level of analysis used in the inferential tests were not individual flights but individual runs. This means, for example, that the permutation test comparing arrival delivery error in the three conditions used a repeated-measure design with only four data points per condition, each one representing the mean (or standard deviation) of a run.

Multiple comparison permutation tests were not performed because they could not have resulted in a p value smaller than .05. More specifically, a two-paired repeated-measure sample design with only four data points in each condition would have only 16 possible permutations and therefore a minimal p value of 1/16 or .06. Following analyses will also not include multiple comparisons when the design will not possibly lead to p values equal or under .05.

Some of the following sections examine the effects of the experimental condition on various metrics or dependent variables. These analyses did not include scenario as an independent variable because scenarios were developed to have similar characteristics and parameters, and were considered equivalent in traffic flow and complexity. Individual analyses suggested that scenario had no effect on the different dependent variables investigated in the result section.

3.1 SYSTEM PERFORMANCE – EN ROUTE AIRSPACE

3.1.1 Arrival Delivery Accuracy[1]

The ability of controllers (and pilots of autonomous aircraft) to meet the STA at the BAMBE meter fix was examined by observing the difference between the STA and the ATA. Controllers and pilots were instructed that delivery within ±15 seconds of the STA was considered on time. The arrival delivery error (ATA-STA) was calculated for all arrival aircraft that reached the meter fix. An aircraft reaching the meter fix after its STA would be indicated by a positive error value. Aircraft ahead of schedule would therefore be indicated by negative error values.

Average arrival delivery error was analyzed to show the location and dispersion of delivery error. The box-and-whiskers plot in Figure 8 presents the median, first and third quartiles, and 10th and 90th percentiles for each condition. The graph depicts a relatively large

1 When a controller resequenced a flight in the arrival schedule, the flight's STA would automatically update ONLY IF the controller used the swap function on the toolbar. Observed differences in absolute arrival delivery error are likely due, in part, to the fact that controllers did not always use the swap function in the Baseline condition. See Section 4.1.1 for a more detailed explanation.

difference in arrival delivery error dispersion between the Baseline and the CE 6 and CE 5 conditions. The greatest range is found in the Baseline condition, with an error range from 559 seconds behind schedule to 494 seconds ahead of schedule. The plot also shows only a small difference in central tendency between conditions, with delivery error evenly distributed around similar central values (i.e., approximately 10-15 seconds behind schedule).

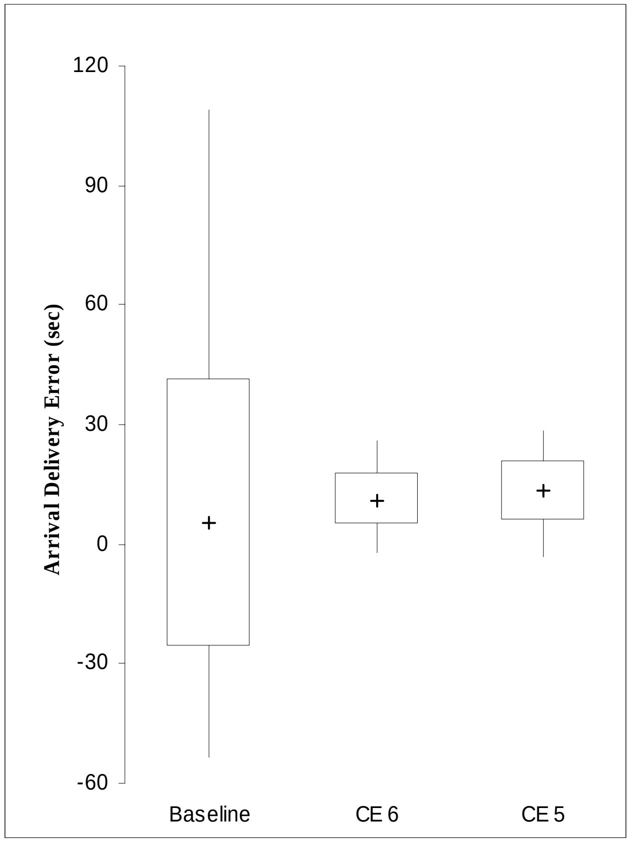

Figure 8. Arrival Delivery Error According to Condition.

Figure 9 and Table 10 presents the average arrival delivery error and standard deviation for the four runs of each of the three conditions. The data indicate that, on average, flights reached the meter fix after the STA in all runs except in the fourth run of the Baseline condition, where flights reached the meter fix 13.8 seconds before the STA. Furthermore, the average arrival delivery error was greater than the ±15-second level of tolerance in only two runs – the first Baseline and the second CE 6. A one-way repeated-measures permutation test showed that means did not differ significantly across conditions ($F[2, 6]= 0.005$, $p=1$).

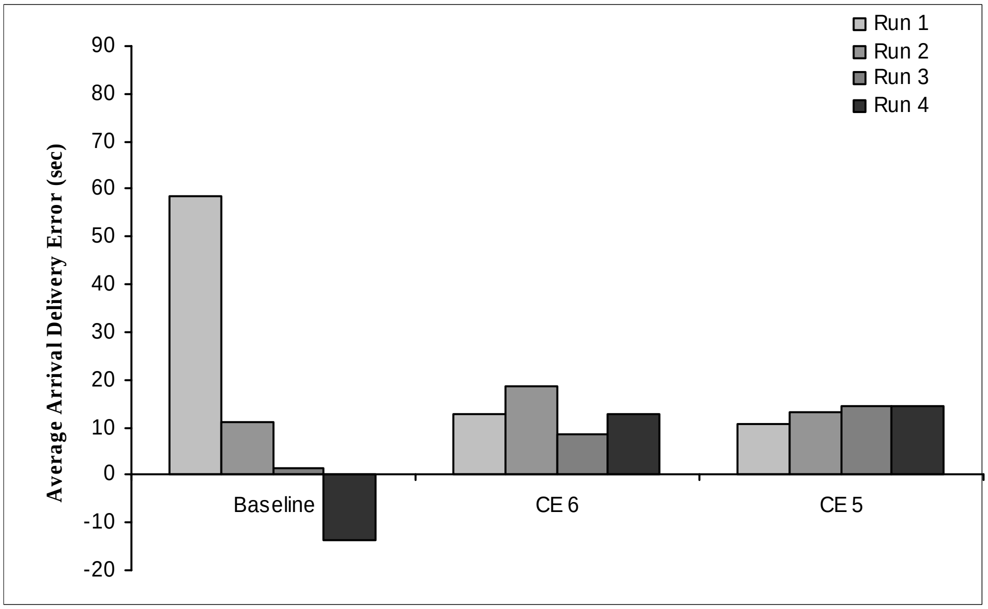

Figure 9. Average Arrival Delivery Error According to Condition and Run.

Table 10. Average Arrival Delivery Error According to Condition and Run.

	Condition					
	Baseline		**CE 6**		**CE 5**	
	M	*SD*	*M*	*SD*	*M*	*SD*
1st Run	54.8	98.8	12.8	13.9	10.6	16.2
2nd Run	11.2	99.2	18.8	47.7	13.3	16.9
3rd Run	1.3	127.1	8.4	19.8	14.4	21.2
4th Run	-13.8	55.2	12.7	37.5	14.2	19.2
Mean	14.3	95.1	13.2	29.7	13.1	18.4
Std. Dev.	31.2	29.7	4.3	15.6	1.7	2.3

One measure of arrival delivery accuracy is given by the dispersion (e.g., variance) of the delivery errors. The larger the variance, the more difficult it was to consistently deliver aircraft on schedule. The dispersion of the delivery errors varied greatly between experimental conditions (Table 10). The average standard deviation was the greatest in the Baseline condition (95.1) and greater in the CE 6 (29.7) than in the CE 5 (18.4) condition. The Baseline run with the smallest standard deviation (55.2 in the fourth run) was higher than the standard deviation of all the runs in the CE 6 and CE 5 conditions. A one-way repeated-measures permutation test showed that standard deviations were significantly different across conditions ($F[2, 6]= 15.49$, $p=.02$).

A second way of evaluating arrival delivery accuracy was to examine the absolute value of the arrival delivery error (|ATA-STA|). Absolute arrival delivery error was analyzed to show the magnitude of the error. The box-and-whiskers plot (median, first and third quartiles, and 10th and 90th percentiles) in Figure 10 depicts a fairly large difference in central tendency between the Baseline and the CE 6 and CE 5 conditions.

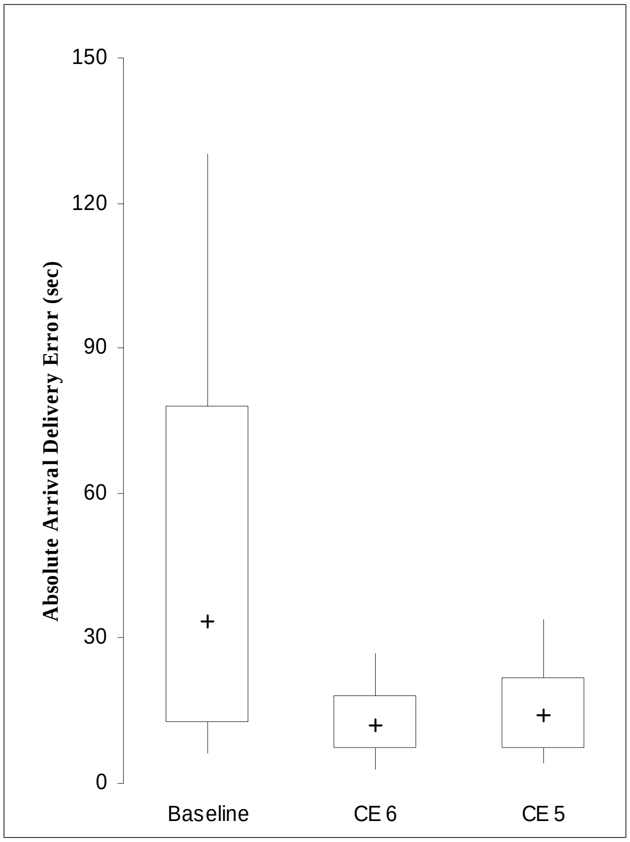

Figure 10. Absolute Arrival Delivery Error According to Condition.

Figure 11 and Table 11 presents the average-absolute-arrival delivery error (and standard deviation) for each run of each condition. The average for the Baseline condition, 59.9, was much greater than the average for the CE 6 and CE 5 conditions (17.5 and 17.1, respectively). A one-way repeated-measures permutation test revealed a significant difference across means ($F[2, 6]=15.50$, $p=.02$).

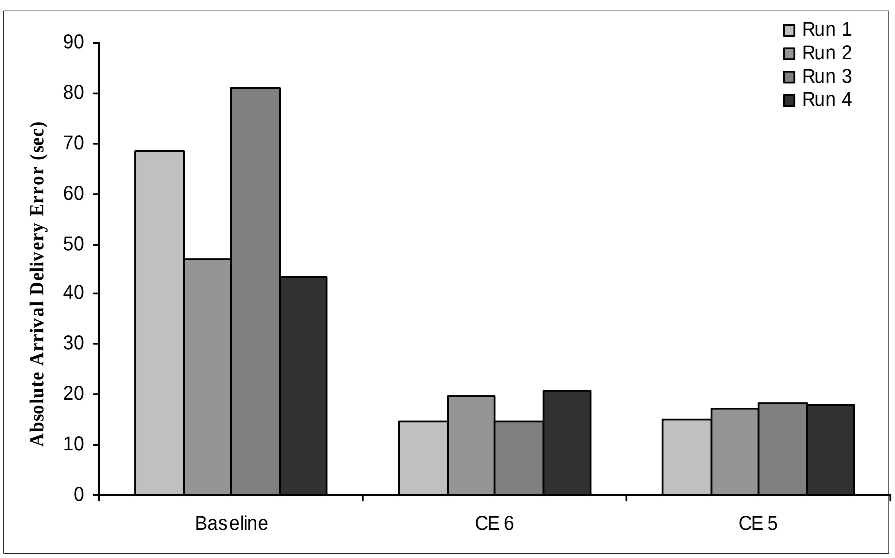

Figure 11. Average Absolute Arrival Delivery Error According to Condition and Run.

Table 11. Average Absolute Arrival Delivery Error According to Condition and Run.

	Condition					
	Baseline		CE 6		CE 5	
	M	*SD*	*M*	*SD*	*M*	*SD*
1st Run	68.3	92.0	14.7	11.9	15.0	12.1
2nd Run	47.0	87.8	19.8	47.3	17.3	12.7
3rd Run	81.2	96.7	14.7	15.5	18.3	17.9
4th Run	43.3	36.2	20.8	33.5	17.8	15.9
Mean	59.9	78.2	17.5	27.1	17.1	14.7
Std. Dev.	18.0	28.2	3.3	16.5	1.5	2.7

The only flights that were equipped to operate in a distributed manner (i.e., having the authority to negotiate a new trajectory and free maneuver) were flown via the PCPlane workstations. PCPlane flights were under ATC control in the Baseline condition. However, PCPlane flights were given authority to negotiate their flight path in the CE 6 condition and to free maneuver in the CE 5 condition.

The following analysis examines the accuracy of PCPlane flight arrival delivery according to condition. Figure 12 presents the mean-absolute-arrival delivery error according to condition and run. Table 12 presents the same information with the addition of standard deviations for each run. The average mean and standard deviations of the four runs were the highest in the Baseline condition. An inferential test concluded that mean delivery error varied significantly across conditions ($F[2,6]=6.51$, $p=.02$ as determined by a randomization test), as it did for the standard deviations ($F[2,6]=5.12$, $p=.009$ as determined by a randomization test).

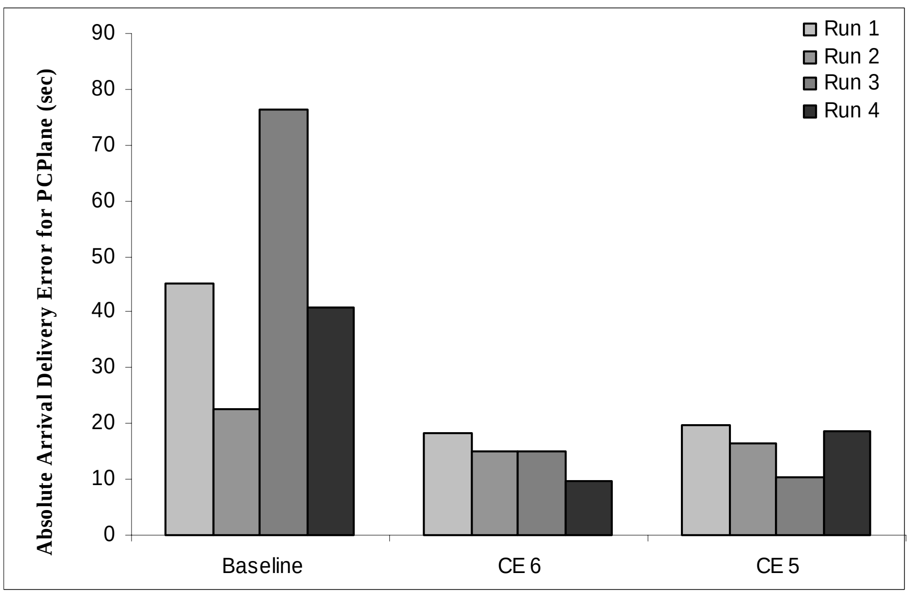

Figure 12. Average Absolute Arrival Delivery Error for PCPlane Flights According to Condition and Run.

Table 12. Average Absolute Arrival Delivery Error for PCPlane Flights According to Condition and Run.

	Condition								
	Baseline			CE 6			CE 5		
	M	*SD*	n	*M*	*SD*	n	*M*	*SD*	n
1st Run	45.2	62.8	5	18.2	11.2	6	19.9	24.0	7
2nd Run	22.5	17.4	4	15.0	7.7	7	16.3	13.0	6
3rd Run	76.3	88.2	7	15.1	7.0	7	10.3	3.8	7
4th Run	41.0	28.3	5	9.9	7.7	7	18.6	11.9	5
Mean	46.2	49.2		14.5	8.4		16.3	13.2	
Std. Dev.	22.3	32.4		3.4	1.9		4.2	8.3	

3.1.2 Arrival Spacing

The time intervals between successive aircraft were recorded (in seconds) as they crossed the meter fix. During the simulation, the TMA was set up to deliver aircraft at the meter fix with a minimum spacing of 7 miles, which equates to an 82-second interval between successive aircraft. As mentioned, controllers and pilots were asked to deliver aircraft within ±15 seconds of the STA, so arrival spacing of less than 68 seconds was considered a deviation from the meter fix crossing minimum. Figure 13 shows the proportion of flights that were spacing below the 68-second minimum. An inferential test revealed that the difference in the proportion of flights was not statistically significant ($F[2,6]=1.13$, $p=.38$ as determined by a permutation test).

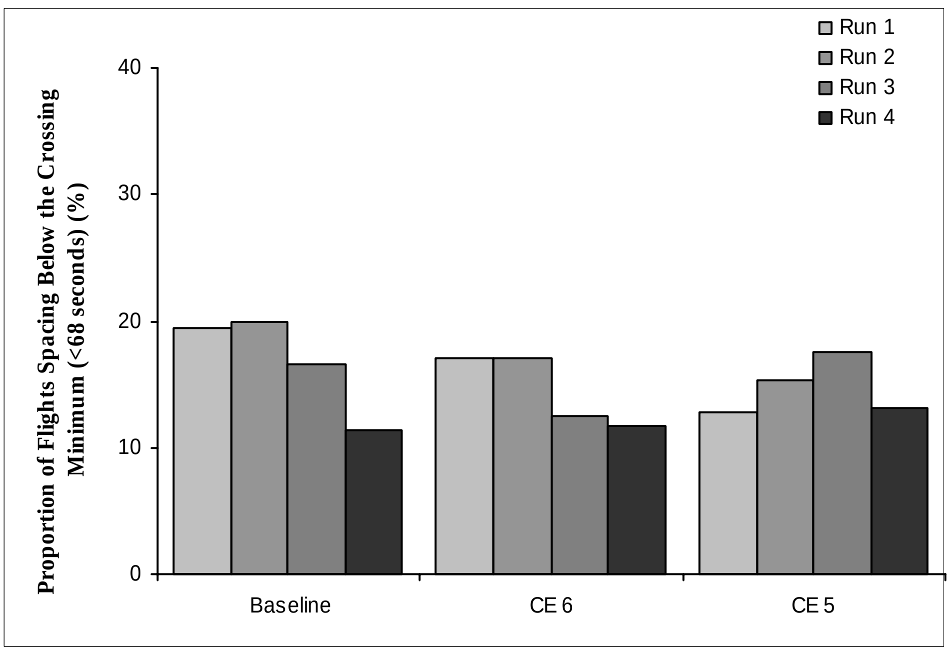

Figure 13. Proportion of Flights Below the 68-Second Crossing Minimum According to Condition and Run.

The box-and-whiskers plot (median, first and third quartiles, and 10th and 90th percentiles), presented in Figure 14, describes the distribution of spacing intervals below the crossing minimum. The plot suggests that spacing intervals deviated the most in the Baseline condition. However, an inferential test comparing the mean spacing deviations, as presented in Table 13, indicated that these spacing deviations did not vary significantly across the three conditions (F[2,6]=0.95, p=.42 as determined by a randomization test).

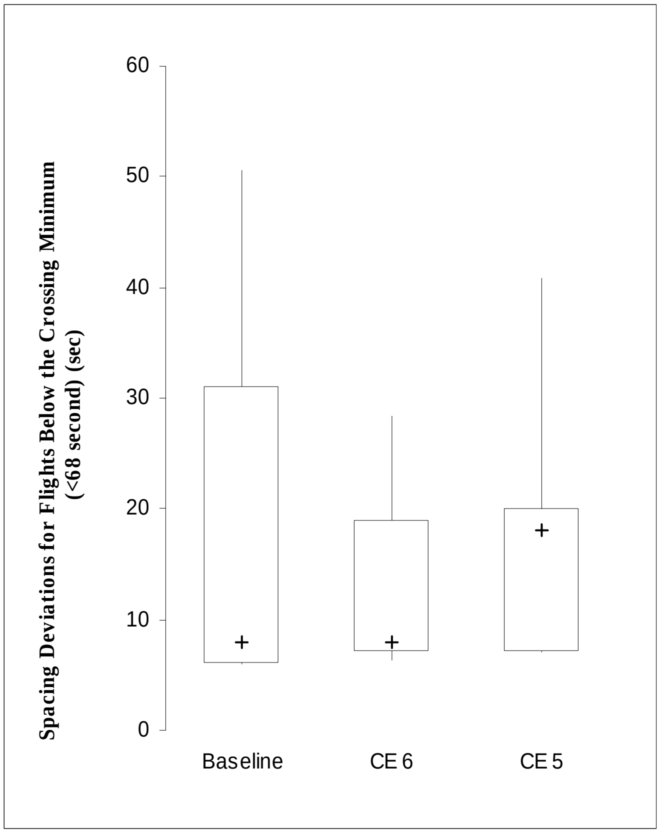

Figure 14. Arrival Spacing Deviations for Flights Below the 68-Second Crossing Minimum According to Condition.

Table 13. Average Arrival Spacing Deviation for Flights Below the 68-Second Crossing Minimum According to Condition and Run.

	Condition											
	Baseline				**CE 6**				**CE 5**			
	M	*SD*	n	%	*M*	*SD*	n	%	*M*	*SD*	n	%
1st Run	22.9	21.8	8	19.5	24.4	22.3	7	17.1	28.8	20.4	5	12.8
2nd Run	27.0	19.3	8	20.0	12.7	9.5	7	17.1	17.5	13.8	6	15.4
3rd Run	23.4	16.6	5	16.7	12.0	6.4	5	12.5	9.0	4.9	7	17.5
4th Run	7.3	1.0	4	11.4	7.0	1.0	5	11.6	19.4	9.2	5	13.2
Mean	20.1	14.7	6.3	16.9	14.0	9.8	6.0	14.6	18.7	12.1	5.8	14.7
Std. Dev.	8.8	9.4	2.1	3.9	7.4	9.0	1.2	2.9	8.1	6.6	1.0	2.2

Although an increase in the number of flights spacing below the 7-mile crossing minimum may reduce system efficiency, these flights do not necessarily represent a safety problem. For the current simulation, the minimum separation requirement at the meter fix was 5 miles, which equates, approximately, to 58 seconds between successive aircraft. Figure 15 shows the proportion of these flights according to condition and run. An inferential test revealed that the difference in the proportion of flights deviating from the separation minimum was not statistically significant (F[2,6]=1.41, p=.27 as determined by a permutation test).

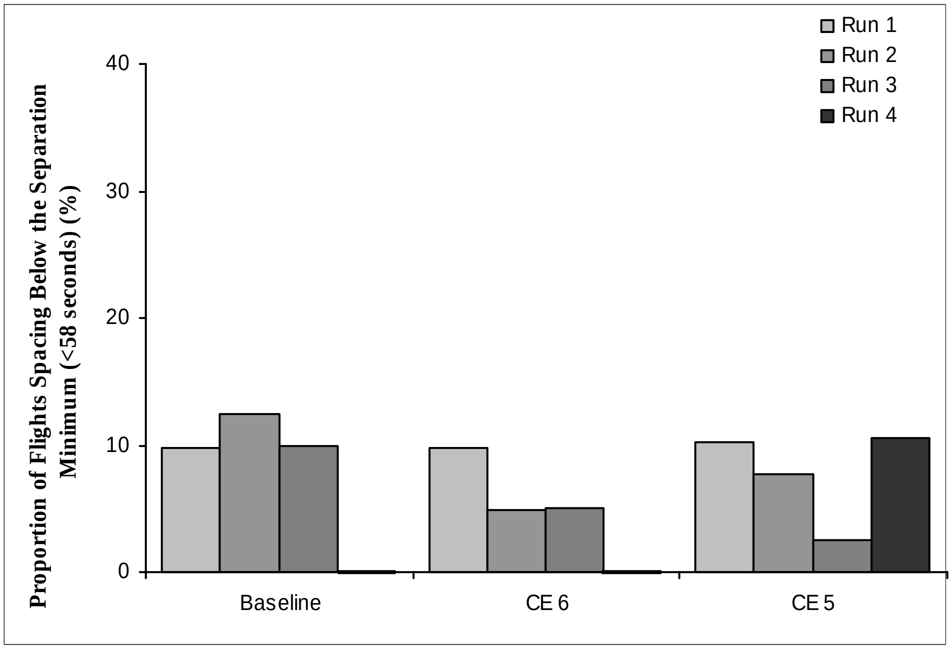

Figure 15. Proportion of Flights Below the 58-Second Separation Minimum According to Condition and Run.

The box-and-whiskers plot (median, first and third quartiles, and 10th and 90th percentiles), presented in Figure 16, describes the distribution of flights that were spacing less than the separation minimum. The plot suggests that spacing deviations for these flights were more severe in the Baseline condition. However, an inferential test comparing the mean spacing deviations for flights below the 58-second minimum, as presented in Table 14, indicated that these deviations did not vary significantly across the three conditions (F[2,6]=0.69, p=.55 as determined by a randomization test).

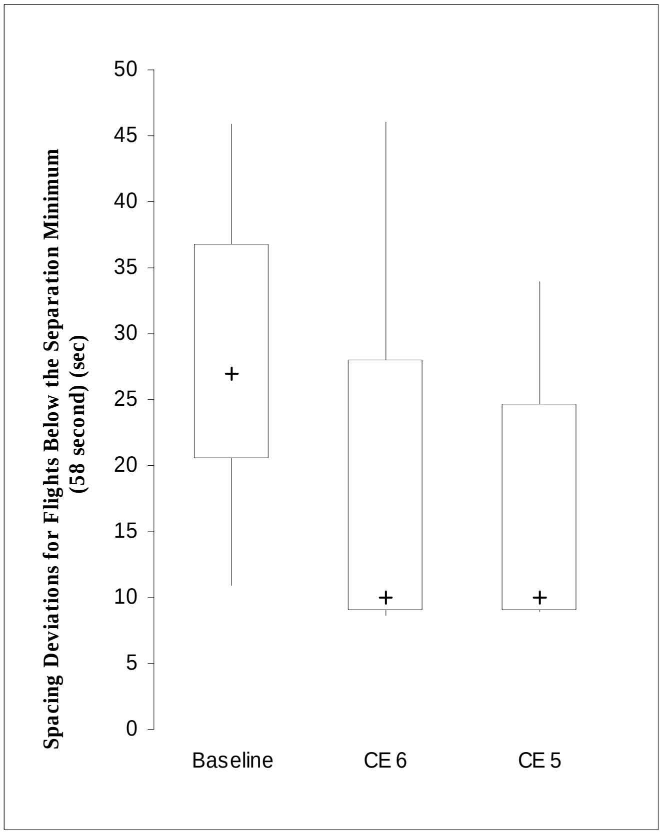

Figure 16. Arrival Spacing Deviations for Flights Below the 58-Second Separation Minimum According to Condition.

Table 14. Average Arrival Spacing Deviation for Flights Below the 58-Second Separation Minimum According to Condition and Run.

	Condition											
	Baseline				CE 6				CE 5			
	M	*SD*	n	%	*M*	*SD*	n	%	*M*	*SD*	n	%
1st Run	29.8	18.6	4	9.8	27.8	21.1	4	9.8	24.3	18.8	4	10.3
2nd Run	28.8	13.7	5	12.5	15.5	9.2	2	4.9	17.3	13.6	3	7.7
3rd Run	25.0	6.9	3	10.0	9.0	1.4	2	5.0	10.0	0.0	1	2.5
4th Run	0.0	0.0	0	0.0	0.0	0.0	0	0.0	12.8	6.2	4	10.5
Mean	20.9	9.8	3.0	8.1	13.1	7.9	2.0	4.9	16.1	9.7	3.0	7.7
Std. Dev.	14.1	8.1	2.2	5.5	11.7	9.7	1.6	4.0	6.2	8.3	1.4	3.7

Note: PCPlane results were not analyzed separately due to the low frequency of flights that deviated from the crossing minimum and separation minimum.

3.1.3 Altitude Deviation

During the simulation, all PCPlane flights were given a meter fix crossing restriction of 11,000 feet, while MACS flights were given 11,100. This restriction included a ±299-foot buffer in which flights were considered on-altitude. Flights outside the buffer were judged to be off-altitude. Figure 17 shows that the proportion of flights that deviated from the assigned altitude was noticeably larger in the Baseline condition than in the CE 6 and CE 5 conditions. A one-way repeated-measures permutation test showed that the proportion of off-altitude flights was significantly different across conditions ($F[2, 6]= 12.35$, $p=.009$).

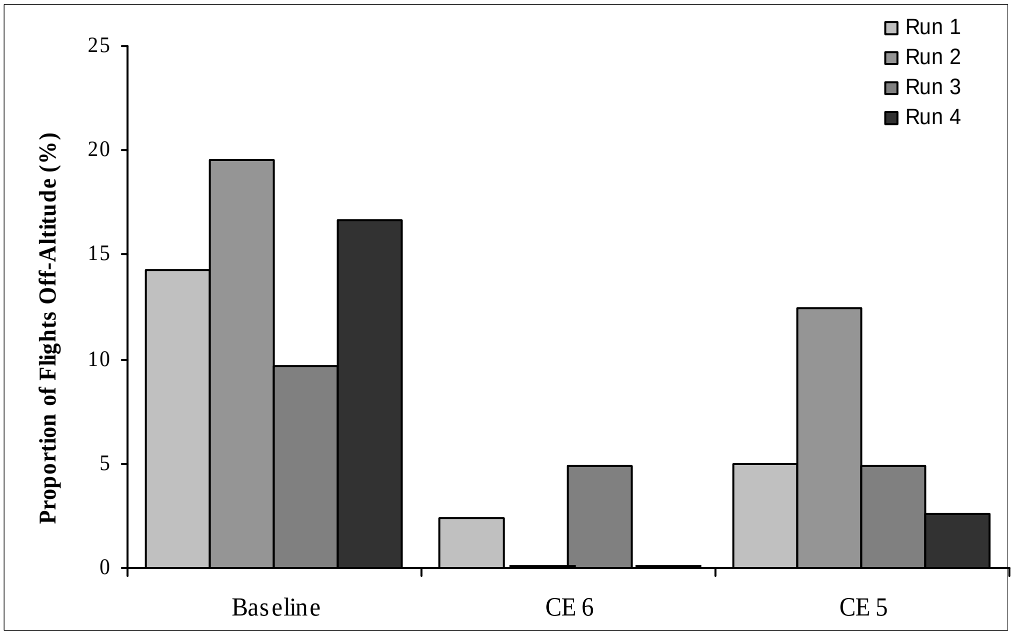

Figure 17. Proportion of Flights Off-Altitude (>±299 ft) at Meter Fix According to Condition and Run.

The box-and-whiskers plot presents the median, first and third quartiles, and 10th and 90th percentiles of off-altitude flights for each condition (Figure 18). The plot shows that the dispersion of altitude deviations was greatest in the CE 5 condition. An inferential test comparing the mean altitude deviations for off-altitude flights, as presented in Table 15, found that the differences were not significant ($F[2,6]=2.75$, $p=.17$ as determined by a randomization test).

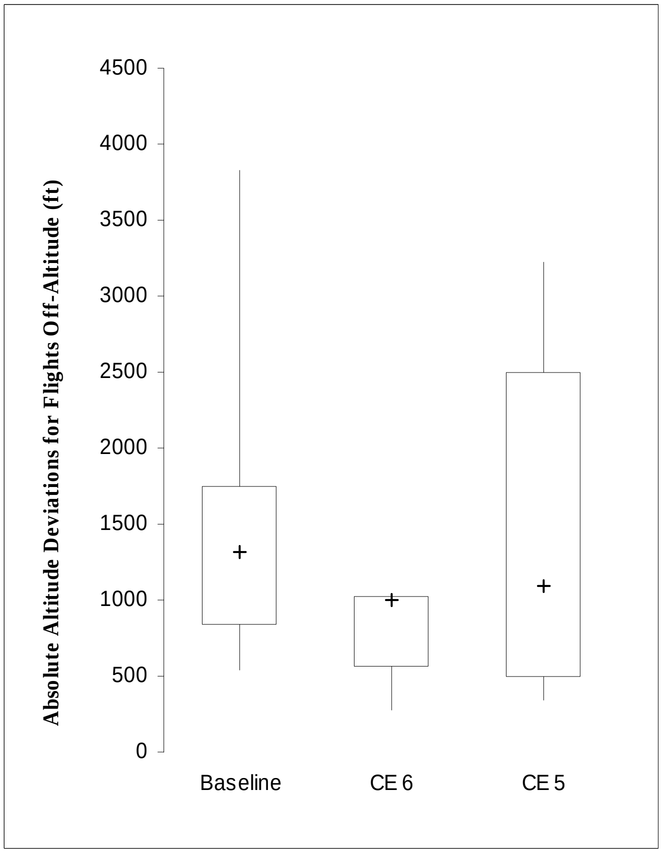

Figure 18. Absolute Altitude Deviation for Flights Off-Altitude (>±299 ft) at Meter Fix According to Condition.

Table 15. Average Absolute Altitude Deviation for Flights Off-Altitude (>±299 ft) at Meter Fix According to Condition and Run.

	Condition											
	Baseline				CE 6				CE 5			
	M	*SD*	n	%	*M*	*SD*	n	%	*M*	*SD*	n	%
1st Run	1245.5	868.9	6	14.3	1050.0	0.0	1	2.4	4245.0	2811.5	2	5.0
2nd Run	1499.0	1079.5	8	19.5	0.0	0.0	0	0.0	1182.8	1061.9	5	12.5
3rd Run	2163.7	1600.3	3	9.7	551.0	632.2	2	4.9	359.5	260.9	2	4.9
4th Run	2084.2	1604.9	6	16.7	0.0	0.0	0	0.0	2583.0	0.0	1	2.6
Mean	1748.1	1288.4	5.8	15.0	400.3	158.0	0.8	1.8	2092.6	1033.6	2.5	6.2
Std. Dev.	447.3	372.9	2.1	4.2	505.1	316.1	1.0	2.3	1703.4	1268.5	1.7	4.3

Note: PCPlane results were not analyzed separately due to the low frequency of flights that deviated from the assigned crossing altitude.

3.1.4 En Route Altitude

During the en route phase, which included the initial arrival descent, the average altitude was calculated at specified distances (i.e., 160 and 60 nm) from the meter fix. The 160-nm distance approximates the top of descent for most flights, while the 60-nm distance is approximately the midpoint of a flight's descent to the terminal boundary. Figure 19 and Figure 20 present the average altitude according to condition and run for the 160 and 60 nm distances, respectively. Figure 19 shows that the average altitude was fairly consistent across conditions at 160 nm from the meter fix. In contrast, at 60 nm from the meter fix, the average altitude appears to be higher in the CE conditions than in the Baseline (Figure 20). Table 16 presents the results of two one-way repeated-measures permutation tests, which confirm that the differences were not significant at the 160 nm distance ($F[2,6]=2.04$, $p=.22$) but were significant at the 60 nm distance ($F[2,6]=55.02$, $p=.04$).

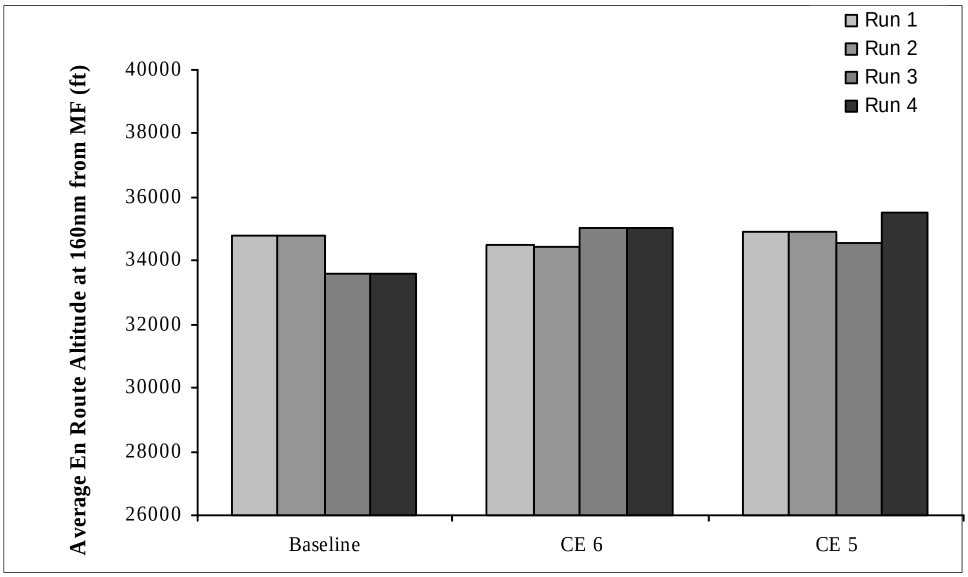

Figure 19. Average En Route Altitude at 160 nm from Meter Fix According to Condition and Run.

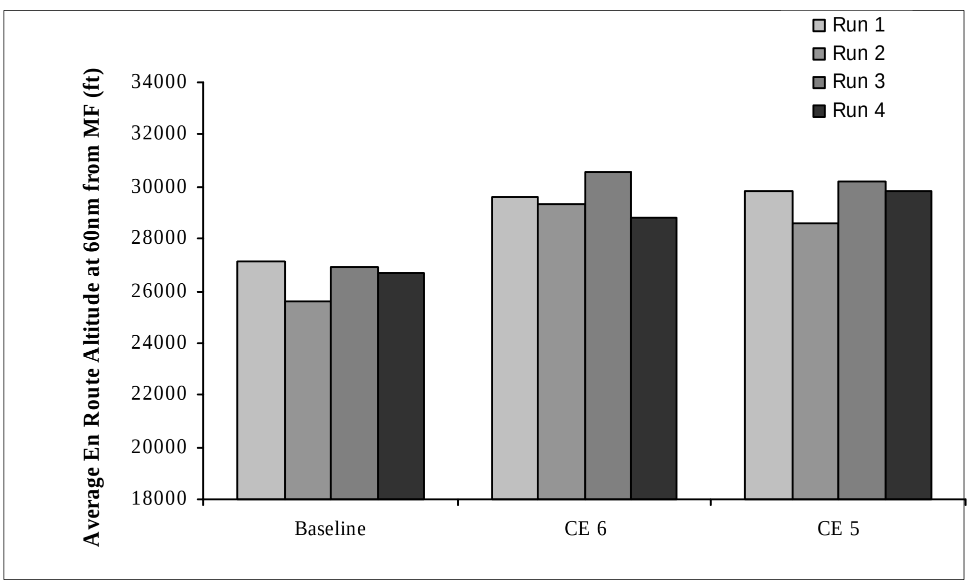

Figure 20. Average En Route Altitude at 60 nm from Meter Fix According to Condition and Run.

Table 16. Average En Route Altitude According to Condition, Run, and Distance from Meter Fix.

Distance from Meter Fix (nm)	Run	Baseline		CE 6		CE 5		F(2,6)	p
		M	*SD*	*M*	*SD*	*M*	*SD*		
160	**1st**	34787	2301	34502	2500	34911	2699	2.04	.22
	2nd	34774	2368	34430	2250	34942	2350		
	3rd	33570	2924	35031	2557	34546	2279		
	4th	33587	3295	35053	2266	35518	2370		
	Mean	34179	2722	34754	2393	34979	2424		
	Std. Dev.	694	473	334	158	402	187		
60	**1st**	27144	3636	29649	2641	29825	2164	55.02	.04*
	2nd	25627	2464	29332	2788	28574	3123		
	3rd	26931	3624	30579	1820	30219	2640		
	4th	26692	2693	28802	2430	29821	2355		
	Mean	26598	3104	29591	2420	29610	2571		
	Std. Dev.	674	614	746	426	715	417		

* Statistically significant as determined by randomization tests (α<=.05).

Figure 21 and Figure 22 show the average en route altitude for PCPlane flights at 160nm and 60nm from the meter fix according to condition. Visual examination of the graphs suggests that altitude varied only slightly across conditions, with the Baseline condition having the lowest average altitude at both distances. Table 17 presents the average altitude, standard deviation, and number of observations according to condition, run, and distance from the meter fix. A permutation test was conducted for each distance to determine if altitude varied significantly between conditions. The results, presented in Table 17, show that differences were statistically significant at the 160nm distance (F[2,6]=6.64, p=.04) but not significant at 60nm from the meter fix (F[2,6]=4.53, p=.07).

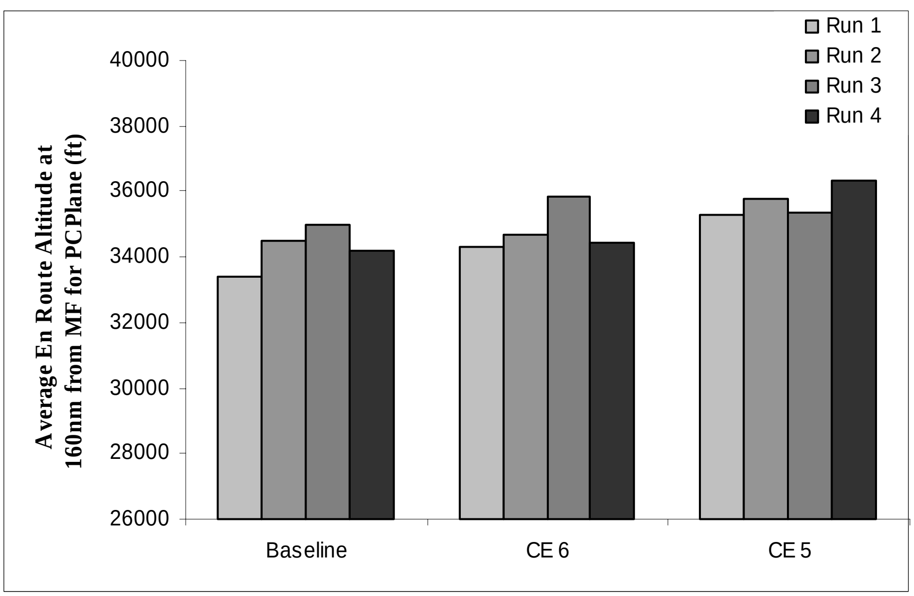

Figure 21. Average En Route Altitude for PCPlane Flights at 160 nm from Meter Fix According to Condition and Run.

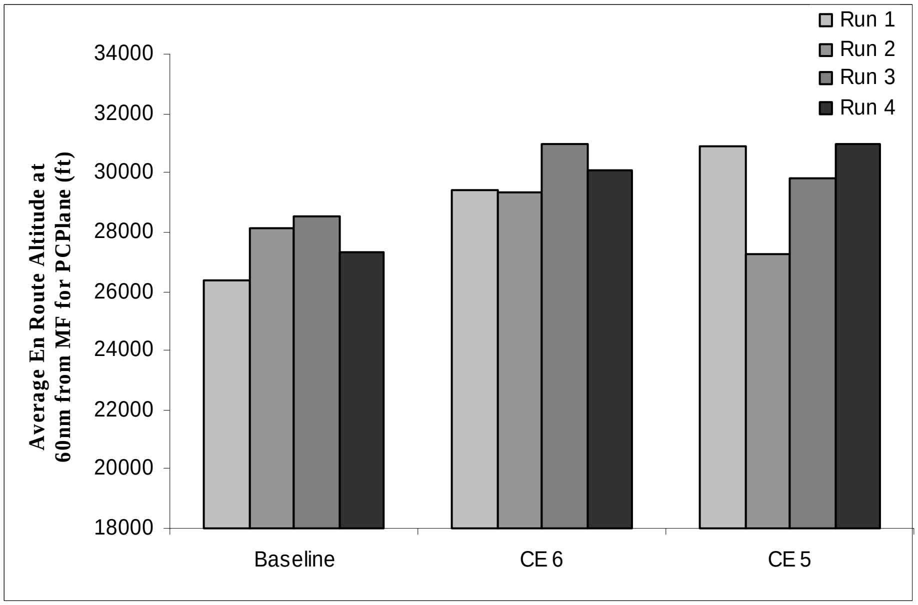

Figure 22. Average En Route Altitude for PCPlane Flights at 60 nm from Meter Fix According to Condition and Run.

Table 17. Average En Route Altitude for PCPlane Flights According to Condition and Distance from Meter Fix.

Distance From Meter Fix (nm)	Run	Baseline			CE 6			CE 5			F(2,6)	p
		M	*SD*	n	*M*	*SD*	n	*M*	*SD*	n		
160	1st	33400	3847	5	34333	3724	6	35285	2928	7	6.64	.04
	2nd	34500	2513	4	34667	1505	6	35803	1784	5		
	3rd	35000	1424	5	35857	1574	7	35333	1506	6		
	4th	34198	3033	5	34429	3207	7	36332	2066	6		
	Mean	34274	2704		34821	2502		35688	2071			
	Std. Dev.	669.9	1014.9		704.6	1132.2		488.4	615.3			
60	1st	26628	4030	5	29494	1296	6	29580	2403	7	4.53	.07
	2nd	26388	3310	4	29398	1196	7	30910	3869	5		
	3rd	28131	4134	7	29336	1434	7	27246	1599	7		
	4th	28548	2778	6	30974	1395	7	29781	1960	7		
	Mean	27424	3563		29801	1330		29379	2458			
	Std. Dev.	1075.6	638.7		785.1	106.8		1537.9	996.5			

3.1.5 Distance Traveled

The distance traveled by each aircraft was recorded starting once it crossed an arc 160 nm from the meter fix. The average distance traveled according to condition and run is presented in Figure 23. Four out of five of the highest distances occurred in Baseline condition runs. Average distance traveled in each condition, shown in Table 18, varied by a few nautical miles. The mean for the Baseline condition (176.5) was highest, followed by CE 6 (169.1) and CE 5 (167.5). Differences between conditions were found to be significant (F[2,6]=4.52, p=.05 as determined by a randomization test).

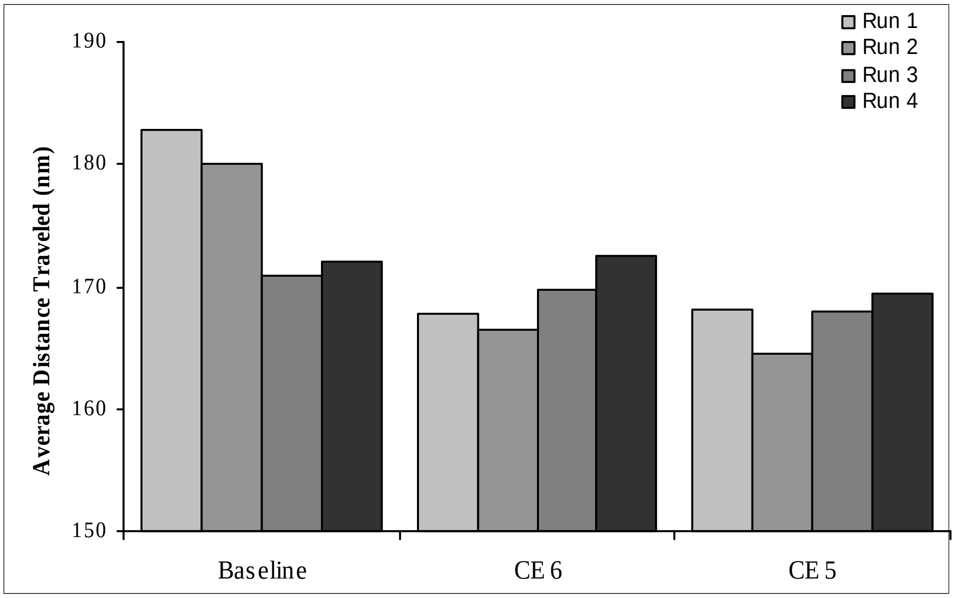

Figure 23. Average Distance Traveled According to Condition and Run.

Table 18. Average Distance Traveled According to Condition and Run.

	Condition					
	Baseline		**CE 6**		**CE 5**	
	M	*SD*	*M*	*SD*	*M*	*SD*
1st Run	182.9	17.3	167.9	6.0	168.2	6.8
2nd Run	180.1	18.1	166.4	5.0	164.6	4.5
3rd Run	170.9	7.9	169.8	9.2	168.0	5.9
4th Run	172.0	10.7	172.5	8.7	169.5	9.9
Mean	176.5	13.5	169.1	7.2	167.5	6.8
Std. Dev.	5.9	5.0	2.6	2.0	2.1	2.3

Figure 24 presents the average distance traveled by PCPlane flights in each simulation run. Table 19 presents the same information with the standard deviation and number of observations for each run. The distance traveled diminished from the Baseline condition (177.0) to CE 6 (169.0), and from CE 6 to CE 5 (164.5); however, differences were not found to be statistically significant (F[2,6]=1.84, p=.17 as determined by a randomization test).

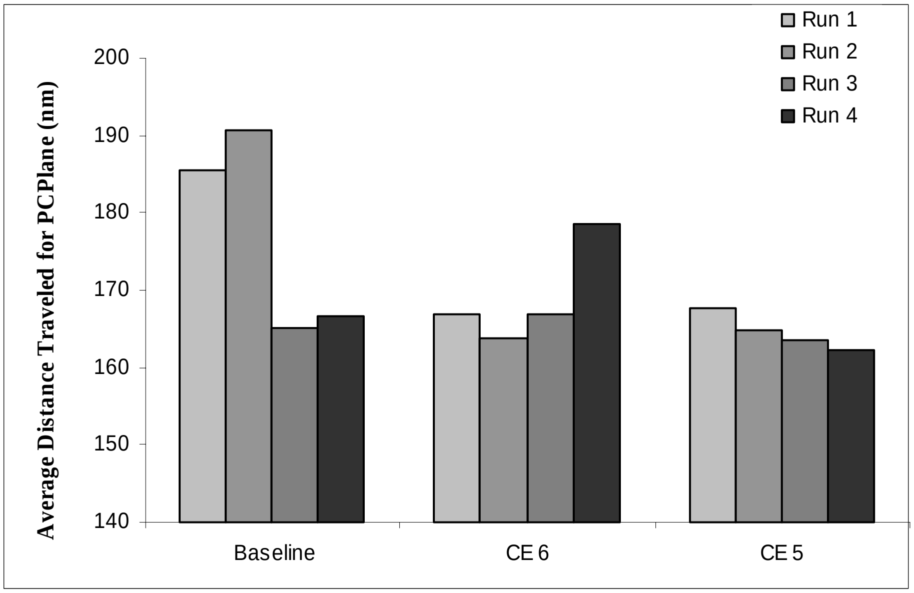

Figure 24. Average Distance Traveled for PCPlane Flights According to Condition and Run.

Table 19. Average Distance Traveled for PCPlane Flights According to Condition and Run.

	Condition								
	Baseline			**CE 6**			**CE 5**		
	M	*SD*	n	*M*	*SD*	n	*M*	*SD*	n
1st Run	185.4	17.7	5	166.8	3.0	6	167.6	8.6	7
2nd Run	190.8	6.8	4	163.8	3.1	6	164.8	7.6	6
3rd Run	165.1	6.2	7	166.9	6.1	7	163.5	1.6	6
4th Run	166.6	4.6	5	178.4	6.0	7	162.3	1.5	7
Mean	177.0	8.8		169.0	4.6		164.5	4.9	
Std. Dev.	13.0	6.0		6.5	1.8		2.3	3.8	

3.1.6 Flight Time

Figure 25 presents the average flight time (in minutes) needed to reach the meter fix after crossing the 160-nm arc. Similarly to the results for distance traveled, four out of five of the highest flight times occurred in Baseline condition runs. Table 20 presents the average flight time to the meter fix according to condition and run. A statistical test was performed to compare the flight times. Results of a randomization test show that the differences were statistically significant ($F[2,6]=6.51$, $p=.05$).

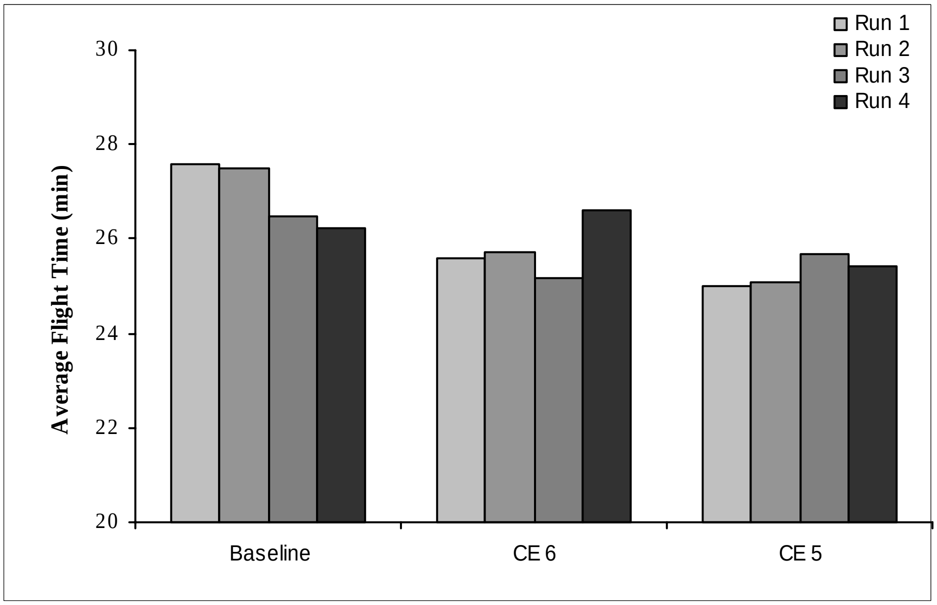

Figure 25. Average Flight Time According to Condition and Run.

Table 20. Average Flight Time According to Condition and Run.

	Condition					
	Baseline		**CE 6**		**CE 5**	
	M	*SD*	*M*	*SD*	*M*	*SD*
1st Run	27.6		25.6		25.0	
2nd Run	27.5		25.7		25.1	
3rd Run	26.5		25.2		25.7	
4th Run	26.2		26.6		25.4	
Mean	26.9		25.8		25.3	
Std. Dev.	0.7		0.6		0.3	

3.1.7 Separation Violations

A separation violation was recorded whenever the separation minima (5 nm laterally and, vertically, ±2,000 feet above FL290 and ±1,000 feet at or below FL290) were violated while both aircraft were in en route airspace. Table 21 shows the count and average number of violations according to condition and run. There was noticeably more separation violations in the Baseline (7) condition than in the CE 6 (1) and CE 5 (1) conditions. However, differences between conditions were found to be only marginally significant ($F[2,6]=6.75$, $p=.07$ as determined by a randomization test).

Table 21. Average Number of Separation Violations According to Condition and Run.

	Condition		
	Baseline	**CE 6**	**CE 5**
	Count	Count	Count
1st Run	3	0	1
2nd Run	1	0	0
3rd Run	1	1	0
4th Run	2	0	0
Total	7	1	1
Mean	1.75	0.25	0.25
Std. Dev.	0.96	0.50	0.50

3.2 System Performance – Terminal Airspace

Unlike en route operations, there was no equivalent baseline condition for terminal operations that could be used in directly evaluating system performance. The analysis presented here is limited to a descriptive account of self-spacing operations as described by the CE 11 operational concept description (Sorensen, 2000). The spacing data were collapsed across CE 6 and CE 5 runs, except where indicated, because spacing operations were identical for the two conditions. The results consist of total 26 spacing trials and represent approximately 40 percent of the total PCPlane/ACFS flights.

3.2.1 Time Spacing

Time spacing was the time (in seconds) the flight was actively involved in self-spacing. It was recorded from the time when the flight deck terminal approach spacing tool was initially turned on until the lead aircraft crossed the FAF or the pilot manually turned off the spacing tool. The average spacing time for PCPlane flights was 275 seconds (*SD* = 113). Interestingly, the difference in time spacing between the CE 6 and CE 5 conditions was almost 1 minute, 244 seconds (*SD* = 102) and 298 seconds (*SD* = 118) respectively (Table 22).

3.2.2 Time to Enter Spacing Assignment

Activating the flight deck terminal approach spacing tool was a four-step process: 1) turning on the tool, 2) setting the lead aircraft, 3) setting the spacing interval, and 4) activating the spacing assignment. The time to enter the spacing assignment (in seconds) was recorded from when the spacing tool was initially turned on until spacing was activated. Results show the average time was 17.2 seconds (*SD* = 4.9). There appears to be a small difference between the CE 6 and CE 5 conditions, 19.1 (*SD* = 4.6) and 15.8 0 seconds (*SD* = 4.8) respectively (Table 22).

3.2.3 Temporal Spacing Assignment

The TRACON controller, at his discretion, assigned a fixed spacing interval that the trailing aircraft was to maintain. The average spacing assignment was 98 seconds (*SD* = 16.8) and did not vary greatly between CE 6 (*M* = 102, *SD* = 15.4) and CE 5 (*M* = 95, *SD* = 17.7) conditions. The spacing intervals assigned ranged from 60 to 120 seconds.

3.2.4 Time from Runway

The elapsed time from when self-spacing began to when the lead aircraft reached the FAF was recorded to determine "where" the TRACON controller initiated spacing. On average, spacing began 258 seconds (*SD* = 115) from the runway. The range was from 78 to 521 seconds. CE 6 and CE 5 conditions showed a noticeable difference, with CE 5 flights (*M* = 297, *SD* = 129) starting spacing approximately 90 seconds further from the runway than CE 6 flights (*M* = 205, *SD* = 68).

3.2.5 Temporal Spacing Change

Temporal spacing change was used to measure the change in the spacing gap (in seconds) and was recorded from the start to the end of self-spacing. A negative value indicates a decrease in the spacing gap. The average change in temporal spacing was 8.2 seconds, with a range from –18.4 to 38.0 seconds. Results for the CE 6 (*M* = 8.4, *SD* = 16.4) and CE 5 (*M* = 8.1, *SD* = 13.4) conditions were not appreciably different (Table 22).

3.2.6 Slant Range

Slant range distance was the metric used to measure changes in spacing distance (in nm). Slant range accounts for changes in vertical and lateral distance, and was recorded from the start to the end of self-spacing. Negative values indicate a decrease in spacing distance. The average slant range was –1.5, with a range from –2.9 to –0.1. The average change was slightly greater in the CE 5 condition (*M* = -1.6, *SD* = 0.7) than in the CE 6 condition (*M* = -1.4, *SD* = 0.9). This was not unexpected because the average time spacing was nearly 1 minute longer in the CE 5 condition than in the CE 6 condition (Table 22).

3.2.7 Rate of Closure

Time-based spacing allowed for spacing compression as aircraft speeds decreased throughout the approach. Rate of closure was the change in lateral spacing distance per minute while self-spacing was on (nm/minute). The average rate of closure was 0.33 (*SD* = 1.8). Differences between the CE 6 (*M* = 0.34, *SD* = 0.23) and CE 5 (*M* = 0.32, *SD* = 0.13) conditions appear to be negligible (Table 22).

3.2.8 Self-Spacing Summary

Table 22 summarizes the results of the in-trail spacing operations that took place within terminal airspace

Table 22. Summary of Self-Spacing Operations.

	Condition					
	Overall		CE 6		CE 5	
	M	*SD*	*M*	*SD*	*M*	*SD*
Total Time Spacing (seconds)	275	113	244	102	298	118
Time to Enter Spacing Assignment into CSD (seconds)	17.2	4.9	19.1	4.6	15.8	4.8
Temporal Spacing Assignment (seconds)	98	16.8	102	15.4	95	17.7
Time from Runway (seconds)	258	115	205	68	297	129
Temporal Spacing Change (seconds)	8.2	14.4	8.4	16.4	8.1	13.4
Change in Slant Range (nm)	-1.5	0.8	-1.4	0.9	-1.6	0.7
Rate of Closure (nm/minute)	0.33	1.8	0.34	0.23	0.32	0.13

3.3 Concept Elements

3.3.1 Trajectory Negotiation

3.3.1.1 Pilot Ratings for Trajectory Negotiation

The post-simulation questionnaire asked several questions concerning the trajectory negotiation concept element. Table 23 presents the ratings derived from the answers to these questions. The first three questions concerned the pilots' overall impression of trajectory negotiation in terms of efficiency, safety, and acceptability. When asked to compare trajectory negotiation with normal current-day operations with respect to efficiency, four pilots rated trajectory negotiation as more efficient, two as the same, and two a less efficient (1 = much less efficient, 5 = much more efficient). The average rating, 3.3 (*SD* = 0.89), suggests that overall pilots thought that trajectory negotiation was comparable to normal current-day operations in terms of efficiency. The next question compared trajectory negotiation and normal current-day operations in terms of safety (1 = much less safe, 5 = much safer). Results show that four pilots found that trajectory negotiation provided a level of safety above normal current-day operations, but an average rating of 3.3 (*SD* = 1.04) suggests that the level of safety was considered comparable. The last question asked pilots to rate the acceptability of trajectory negotiation (1 = unacceptable, 5 = acceptable). The average rating was 3.6 (*SD* = 0.74). All pilots rated the acceptability of the concept element with a score of 3 or higher.

When asked if trajectory negotiation allowed them to implement more favorable routes, pilots' average response was 3.3 (*SD* = 0.71), with seven out of eight pilots giving the concept a rating of at least 3 (1 = not at all, 5 = very much So). Also when asked, pilots felt that controllers were fair during the trajectory negotiation process (1 = not at all, 5 = very much So). Four pilots had a neutral opinion, three thought that controllers were at least fair, and one thought controllers had been very unfair (*M* = 3.3, *SD* = 1.16).

Table 23. Pilot Ratings for Trajectory Negotiation.

Question **(Scale: 1=Negative, 5=Positive)**	**Mean**	**Std Dev.**	**N**
What is your overall impression of trajectory negotiation compared to normal current-day operations in terms of efficiency (e.g., scheduling, and fuel)?	3.3	0.89	8
What is your overall impression of trajectory negotiation compared to normal current-day operations in terms of safety?	3.3	1.04	8
What is your overall impression of trajectory negotiation in terms of acceptability?	3.6	0.74	8
Did the authority to negotiate trajectory changes allow you to implement more favorable routes?	3.3	0.71	8
Do you feel that ATC was fair in the "negotiating" process (i.e., accepting or rejecting the trajectory change requests)?	3.3	1.16	8

3.3.1.2 En Route Controller Ratings for Trajectory Negotiation

Similarly, a post-simulation questionnaire asked en route controllers questions relating to the trajectory negotiation concept element. Results are shown in Table 24. The four en route controllers compared the trajectory negotiation concept in terms of efficiency to normal current-day operations (1 = much less efficient, 5 = much more efficient). The average rating, 4.0 (*SD* = 1.41), suggests that controllers perceived trajectory negotiation operations to be more efficient than normal current-day operations. Likewise, controllers compared their perceived level of safety to normal current-day operations (1 = much less safe, 5 = much safer). Controllers were split on the level of safety trajectory negotiation provided, with two controllers rating it less safe and two rating it safer. The average rating was 3.0 (*SD* = 1.15).

The following four questions relate to the controllers' acceptability of the trajectory negotiation concept. Controllers' average rating for overall acceptability of the trajectory negotiation concept as compared to normal current-day operations (1 = much less acceptable, 5 = much more acceptable) was 3.8 (*SD* = 1.26). This suggests that they generally found the concept acceptable; however, one controller rated it as less acceptable than normal current-day operations. More specifically, controllers were asked to rate the acceptability of three concept-related tasks: pilots modifying speeds without ATC coordination, pilots sending clearance requests, and pilots down linking clearance requests without verbal coordination (1 = completely unacceptable, 5 = completely acceptable). Results suggest that controllers generally found each concept-related task acceptable.

Table 24. En Route Controller Ratings for Trajectory Negotiation.

Question (Scale: 1=Negative, 5=Positive)	Mean	Std Dev.	N
What is your overall impression of the trajectory negotiation operational concept compared to normal current-day operations in terms of efficiency (e.g., scheduling, and fuel)?	4.0	1.41	4
What is your overall impression of the trajectory negotiation operational concept compared to normal current-day operations in terms of safety?	3.0	1.15	4
What is your overall impression of the trajectory negotiation operational concept for handling arrival traffic in terms of acceptability?	3.8	1.26	4
How acceptable was the concept of pilots modifying speeds without ATC coordination?	4.0	1.41	4
How acceptable was the concept of pilots sending clearance requests?	4.8	0.50	4
Were down linked clearance requests with no verbal communications operationally acceptable?	4.7	0.58	3

3.3.2 Free Maneuvering

3.3.2.1 Pilot Ratings for Free Maneuvering

Several items in the post-simulation questionnaire focused on various aspects of the free maneuvering concept. Table 25 presents the results of these questions. When asked about efficiency, ratings show that seven pilots perceived free maneuvering to be more efficient than normal current-day operations; and the eighth pilot thought that free maneuvering was as efficient (1 = much less efficient, 5 = much more efficient). The average rating was 4.0 (*SD* = 0.53). Five pilots perceived free maneuvering to provide a level of safety above normal current-day operations, two found it offered the same level, and one less. The average rating of 3.6 (*SD* = 0.92) suggests that overall free maneuvering was perceived to have provided a level of safety at least comparable to normal current-day operations (1 = much less safe, 5 = much safer). Similarly, the level of safety (1 = much less safe, 5 = much safer) and ease-of-use (1 = much more difficult, 5 = much easier) associated with the rules-of-the-road were perceived to be slightly above that of normal current-day operations. Ratings regarding the acceptability of free maneuvering show that pilots positively rated the concept (*M* = 4.0, *SD* = 0.93) with seven out of eight giving it a rating of 4 or above (1 = unacceptable, 5 = acceptable). The last question focused on the effectiveness of the concept. A majority of pilots considered that free maneuvering allowed them to implement more favorable routes (1 = not at all, 5 = very much So). The average rating was 3.6 (*SD* = 0.74).

Table 25. Pilot Ratings for Free Maneuvering.

Question (Scale: 1=Negative, 5=Positive)	Mean	Std Dev.	N
What is your overall impression of free maneuvering compared to normal current-day operations in terms of efficiency (e.g., scheduling, and fuel)?	4.0	0.53	8
What is your overall impression of free maneuvering compared to normal current-day operations in terms of safety?	3.6	0.92	8
Did the rules-of-the-road used result in a level of safety that was higher/lower than compared to normal current-day operations?	3.6	0.51	8
How easy was it to follow the rules-of-the-road?	3.3	0.59	8
What is your overall impression of free maneuvering in terms of acceptability?	4.0	0.93	8
Did autonomous control allow you to implement more favorable routes during the en route phase?	3.6	0.74	8

The data in the following three tables, Table 26, Table 27, and Table 28, are a compilation of pilot responses to the post-simulation questionnaire. Pilots were asked to specify the reasons for initiating flight path changes other than for meeting their STA and resolving traffic conflicts (Table 26). Reasons included simplifying the flight plan, shortening route and flying time, and increasing aircraft efficiency. Four pilots responded that they did not have any reason for initiating flight path changes except for meeting their STA and resolving traffic conflicts.

Table 26. Pilot Reasons for Initiating Flight Path Changes During Autonomous Operations.

Reason	n
None.	4
To simplify flight plan.	2
For aircraft efficiency.	2
To shorten route and flying time before reaching the freeze horizon.	1
To increase separation traffic ahead or behind.	1

Pilots identified some of the difficulties they encountered in trying to detect and resolve traffic conflicts while under autonomous control (Table 27). Reasons include pilots' lack of familiarity with the tools and rules-of-the-road and not knowing whether an aircraft was autonomous. Some interface shortcomings were also identified.

Table 27. Difficulties Encountered by Pilots Detecting and/or Resolving Conflicts During Autonomous Operations.

Difficulty	n
How to resolve conflicts while meeting PTA. Do you arrive early or late if unable to make time *and* resolve conflict?	1
Selected speed, vertical speed, and altitude changes are not probed by conflict alerting system.	1
When both pilots involved in a traffic conflict attempt to resolve the conflict, confusion and workload increase.	1
Since there is no direct communications between conflicting aircraft, it is hard to figure out each other's intentions.	1
Not being familiar enough with all available tools.	1
Did not always know if the conflict aircraft was managed or autonomous.	1
When conflicts occurred, the CSD traffic display was sometimes too cluttered to start resolution early.	1
Was not always sure which rule-of-the-road applied (needed more training on rules-of-the-road).	1
Unable to make route changes inside the freeze horizon.	1
Some confusion as to ATC assigned speeds and pilots ability to change them.	1

Pilots were also asked to identify concerns they had with respect to roles and responsibilities during autonomous control. As listed in Table 28, pilots reported that workload and heads-down time increased. Pilots also mentioned that not knowing other flights' RTAs and not seeing the big picture made it more difficult for them to make decision that would benefit everyone.

Table 28. Pilot Concerns Regarding Roles and Responsibilities During Autonomous Control.

Concern	n
Everyone (pilots and controllers) has to be on the same page. No other agenda.	1
Definite increase in pilot workload that increases as you near the metering fix.	1
Pilots need the ability to communicate with other aircraft to find out their situation (i.e., RTA and aircraft type for performance capabilities).	1
With the rules given, no problem.	1
Some concern as to seeing the big picture as to the overall flow of traffic and making decisions that work best for everyone.	1
Having heads down during critical flight phases and high workload (e.g., running checklists, briefing approaches, and making changes using RAT when also using weather avoidance radar).	1

3.3.2.2 En Route Controller Ratings for Free Maneuvering

The four en route controllers rated their impressions of several aspects of the free maneuvering concept with respect to normal current-day operations (Table 29). When asked how efficient free maneuvering operations were (1 = much less efficient, 5 = much more efficient), the average rating was 4.3 (*SD* = 0.50). This suggests that controllers perceived free maneuvering to be more efficient than normal current-day operations. As for safety (1 = much less safe, 5 = much safer), controller ratings were generally less favorable for the perceived level of safety that free maneuvering provided. The average rating was 2.5 (*SD* = 0.58). For acceptability (1 = much less acceptable, 5 = much more acceptable), the average rating was 2.5 (*SD* = 0.58). Two controllers found the free maneuvering concept less acceptable than normal current-day operations, while the two others expressed a neutral view.

Controllers were also asked to rate the acceptability of specific aspects of the free maneuvering concept: procedures and phraseology for applying the rules-of-the-road, criteria for canceling autonomous control, and procedures and phraseology for canceling autonomous control (1 = completely unacceptable, 5 = completely acceptable). Table 29 shows that controllers rated these three aspects of free maneuvering above the neutral point.

One aspect of the free maneuvering concept was to employ the low-altitude sector controller (i.e., Bowie) as a "gate keeper" for the entry of arrival aircraft into the TRACON airspace. The controller could deny entry if speed, altitude, or ATA was deemed unacceptable. If so, the aircraft would be vectored for resequencing into the arrival stream. When asked if this role was acceptable for the lower altitude controller, the response was generally neutral or above. The average rating was 3.8 (*SD* = 0.58).

Table 29. En Route Controller Ratings for Free Maneuvering.

Question **(Scale: 1=Negative, 5=Positive)**	**Mean**	**Std Dev.**	**N**
What is your overall impression of the free maneuvering operational concept compared to normal current-day operations in terms of efficiency (e.g., scheduling, and fuel)?	4.3	0.50	4
What is your overall impression of the free maneuvering operational concept compared to normal current-day operations in terms of safety?	2.5	0.58	4
What is your overall impression of the free maneuvering operational concept for handling arrival traffic in terms of acceptability?	2.5	0.58	4
How acceptable were the procedures and phraseology for applying the rules-of-the-road?	3.5	1.29	4
How acceptable were the criteria for autonomous control cancellation?	3.8	1.15	3
How acceptable were the procedures and phraseology for cancellation of autonomous control?	4.0	1.00	3
How acceptable was the concept of having the Bowie controller act as "gate keeper" for entry to controlled (TRACON) airspace?	3.8	0.58	3

Table 30 lists controller responses when asked what prompted them to cancel autonomous control. The three responses were the aircraft's inability to make the RTA, when assistance for conflict resolution was requested, and when aircraft equipment failed.

Table 30. Controller Reasons for Canceling Autonomous Control.

Reason	**n**

Inability to make RTA.	2
When assistance for conflict resolution was requested.	1
When aircraft equipment failed.	1

3.3.3 Terminal Approach Spacing

3.3.3.1 Pilot Ratings for Terminal Approach Spacing

Similarly, pilots were queried on several aspects of the terminal approach self-spacing concept. Results are presented in Table 31. Pilots' overall impression of terminal approach spacing in terms of efficiency was above that of normal current-day operations. Only one pilot thought the level of efficiency provided by terminal approach spacing was the same as normal current-day operations, while the seven other pilots rated it above (1 = much less efficient, 5 = much more efficient). The average rating was 4.3 (*SD* = 0.71). As for safety, four pilots found terminal approach spacing to provide a level of safety above that of normal current-day operations (1 = much less safe, 5 = much safer). The four other pilots judged it to be the same. Pilots gave an average safety rating of 3.9 (*SD* = 0.99) to this concept element. When asked about the acceptability of the concept, six out of eight pilots rated acceptability above the neutral point (1 = unacceptable, 5 = acceptable). The average rating was 4.0 (*SD* = 0.76).

When they had entered the TRACON airspace, pilots monitored the frequency for a self-spacing clearance, which consisted of a lead aircraft and spacing assignment. The spacing interval ranged from 60-120 seconds, with an average of 98 seconds. Five out of eight pilots felt the self-spacing intervals assigned by controllers were acceptable. The three other pilots gave them a neutral rating of 3 (1 = unacceptable, 5 = acceptable). The average rating was 3.9 (*SD* = 0.83).

Table 31. Pilot Ratings for Terminal Approach Spacing.

Question **(Scale: 1=Negative, 5=Positive)**	**Mean**	**Std Dev.**	**N**
What is your overall impression of terminal self-spacing compared to normal current-day operations in terms of efficiency (e.g., scheduling, and fuel)?	4.3	0.71	8
What is your overall impression of terminal self-spacing compared to normal current-day operations in terms of safety?	3.9	0.99	8
What is your overall impression of terminal self-spacing in terms of acceptability?	4.0	0.76	8
Did you feel that the self-spacing intervals assigned to your aircraft were acceptable?	3.9	0.83	8

Table 32 shows pilot suggestions for a better implementation of self-spacing as reported in the post-simulation questionnaire. Two pilots suggested that the ability to specify a minimum distance with the spacing tool be implemented. Two pilots suggested that the spacing tool be distance instead of time based (or that the possibility be examined). Another pilot reported that the method for entering spacing parameters should be improved.

Table 32. Pilot Suggestions Regarding Terminal Approach Spacing Implementation.

Suggestion	n
Implement minimum distance spacing in PDA-based approaches.	2
Interval should be based on distance and not time.	2
Easier method of entering spacing parameters, involving fewer steps.	1
TRACON should not change spacing on PCPlane after intercepting LOC.	1

3.3.3.2 TRACON Controller Ratings for Terminal Approach Spacing

The interpretation of the TRACON controller data is limited. To begin with, data collected at the TRACON controller position is from only one participant. This alone severely limits the conclusions and generalizations that can be derived from these results. Secondly, the controller developed his own style of implementing the terminal approach spacing task and used the tools in a manner that differed from the way he was taught during training. It was observed that the controller continued to learn and change his strategy throughout the simulation. Therefore, the results presented below may reflect the general difficulties encountered rather than specific aspects of the concept.

The controller compared the terminal approach spacing concept to normal current-day operations in terms of efficiency (1 = much less efficient, 5 = much more efficient), safety (1 = much less safe, 5 = much safer), and acceptability (1 = unacceptable, 5 = acceptable) (Table 33). For all three attributes, the controller ratings were below that of normal current-day operations. In addition, when asked to rate workload levels associated with in-trail spacing operations, he consistently rated the level of workload as unfavorable.

Table 33. TRACON Controller Ratings for Terminal Approach Spacing.

Question (Scale: 1=Negative, 5=Positive)	Rating	Std Dev.	N
What is your overall impression of the in-trail spacing operational concept compared to normal current-day operations in terms of efficiency?	2	-	1
What is your overall impression of the in-trail spacing operational concept compared to normal current-day operations in terms of safety?	2	-	1
What is your overall impression of the in-trail spacing operational concept for handling arrival traffic in terms of acceptability?	2	-	1
How did the workload in in-trail spacing condition compare to Baseline?	1	-	1
How difficult was it to monitor and maintain separation during in-trail spacing operations?	2	-	1
How difficult was it to deliver aircraft on schedule during in-trail spacing operations?	2	-	1

Table 34 lists the comments and concerns the TRACON controller expressed about the terminal approach spacing concept and in-trail spacing procedure.

Table 34. TRACON Controller Comments About the Terminal Approach Spacing Concept.

Comment	n
Unclear when self-spacing should be applied.	1
Any speeds slower than the expected speeds caused severe problems for the pilots.	1
Pilots had a difficult time reacting in a timely way to maintain spacing.	1
Increased workload over today's operation. However, familiarity with the operation and standard in trail-time-spacing application rules may reduce it.	1
Unless closely monitored, efficiency and capacity suffer.	1

3.4 HUMAN PERFORMANCE

3.4.1 Pilot Workload

Post-run questionnaire data show that pilot ratings of mental workload (i.e., how much mental and perceptual activity was required, for example, thinking, deciding, calculating, remembering, looking, searching) and temporal demand (i.e., how much time pressure did the pilot feel due to the rate or pace at which tasks or task elements occurred) were perceived to be below that of normal current-day operations for all three phases of flight (1 = much lower, 5 = much higher). Figure 26 and Figure 27 show these results, respectively.

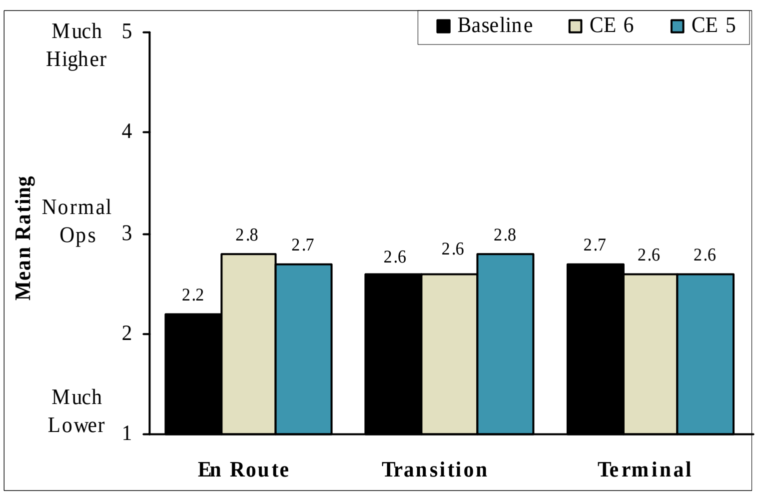

Figure 26. Pilot Comparison Rating of Mental Workload According to Condition.

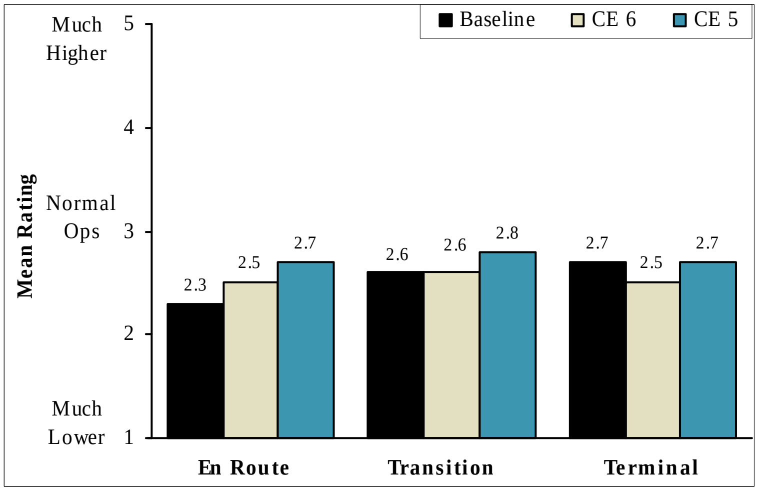

Figure 27. Pilot Comparison Rating of Temporal Demand According to Condition.

Pilots rated their level of confusion (e.g., problems understanding procedures or tool usage) after each run with respect to each phase of flight (i.e., en route, transition, and terminal) (1 = much lower, 5 = much higher). The average level of confusion rating was below that of normal current-day operations (Figure 28). The highest rating of confusion for each phase of flight occurred in the CE 5 condition.

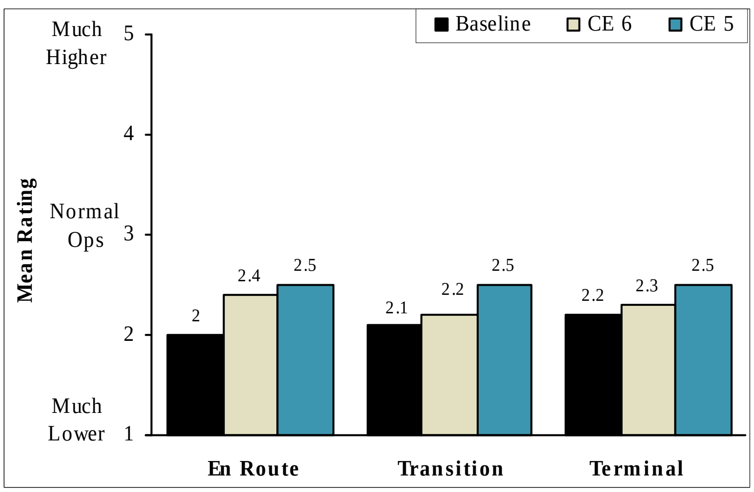

Figure 28. Pilot Comparison Rating of Confusion According to Condition.

3.4.2 Pilot Situation Awareness

Pilots' average ratings for situation awareness (SA) indicated that SA was perceived to be above that of normal current-day operations. (Situation awareness, which is commonly referred to as "staying ahead of the aircraft," involves processing the relevant information to develop a thorough understanding of the current situation, which facilitates taking the appropriate actions in a timely manner.) Figure 29 shows the average situation awareness ratings for each phase of flight (1 = much lower, 5 = much higher). The graphs reveal that SA ratings did not vary across conditions.

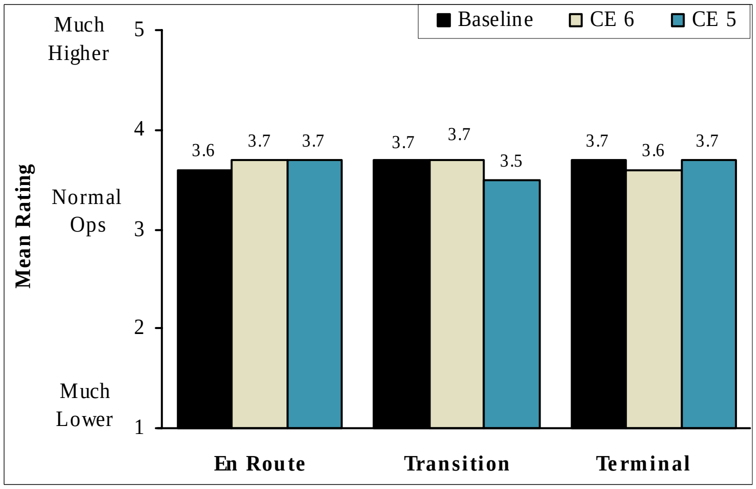

Figure 29. Pilot Comparison Rating of Situation Awareness According to Condition.

3.4.3 En Route Controller Performance

En route controllers rated their performance on several specific tasks, as well as overall performance, after each run. Results are shown in Figure 30. When asked to rate their overall performance (i.e., how successful they were in accomplishing the goals of the task and how satisfied they were with their performance in accomplishing these goals) (1 = Low, 5 = High), controllers gave average ratings that were above the neutral point (i.e., 3). The lowest performance rating was for the Baseline condition. En route controllers also rated the ease with which they could monitor and maintain separation (1 = very difficult, 5 = very easy). The Baseline, CE 5, and CE 6 conditions all received average ratings above the neutral point, with the CE 5 condition receiving the highest rating. For perceived difficulty in maintaining the arrival schedule (1 = difficult, 5 = easy), the average ratings suggest that controllers had the easiest time maintaining the arrival schedule in the CE 6 condition (4.2) and the most difficult time in the Baseline condition (2.1). Average ratings for success in maintaining safety in their sector (1 = not successful, 5 = successful) indicate that all conditions were above the neutral point.

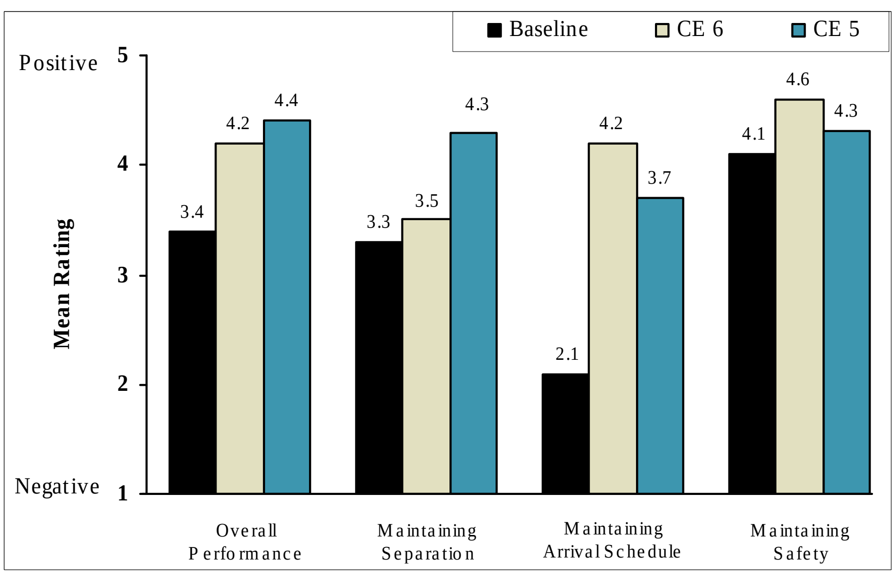

Figure 30. En Route Controller Rating of Performance According to Condition.

3.4.4 En Route Controller Workload

Figure 31 shows the average ATWIT rating by condition and position (1 = Low, 7 = High). Controller workload ratings were higher in the Baseline condition than in the CE 6 and CE 5 conditions for all four positions. A one-way repeated-measures permutation test comparing the run averages for each condition showed that differences in perceived workload were statistically significant only for the Amarillo position (F=11.94, p=.03 as determined by a randomization test). Figure 32 shows controller mean ATWIT ratings distributed over the length of the run. The value of a specific time interval is the average rating for the four runs within each condition. The graph suggests that on average perceived workload was slightly higher in the Baseline condition after 24 minutes from the start of the run.

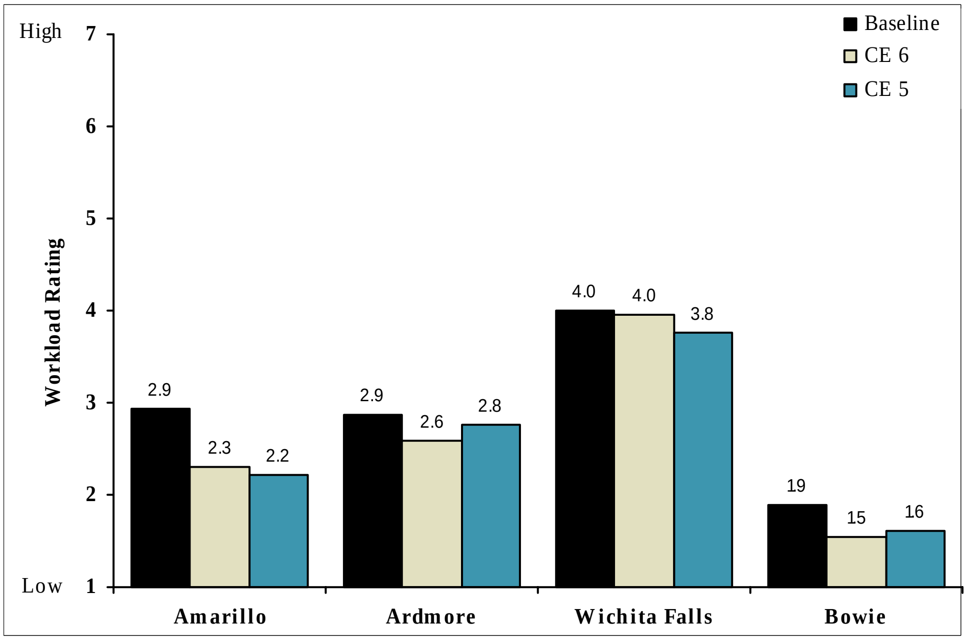

Figure 31. En Route Controller ATWIT Workload Rating According to Condition and Position.

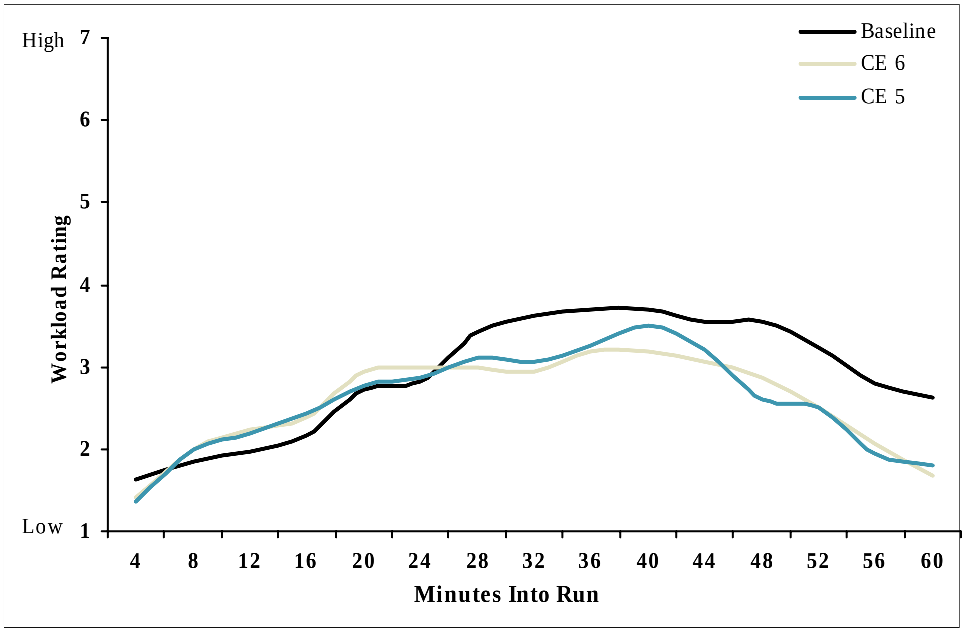

Figure 32. En Route Controller Mean ATWIT Workload Rating According to Condition and Minutes into the Run.

En route controllers rated the acceptability of the workload and made workload comparisons to normal current-day operations and the Baseline condition (Figure 33). In the Baseline condition, controllers rated workload as compared to normal current-day operations, while the CE 6 and CE 5 conditions were compared to the Baseline Condition (1 = Greatly Increased, 5 = Greatly Decreased). The controllers indicated that their perceived workload during the Baseline condition was comparable to normal current-day operations. Average workload scores for the CE 6 and CE 5 conditions indicate that controllers' workload was perceived to be less than that of the Baseline condition, 4.5 and 4.3, respectively. Controllers also rated the acceptability of the workload for each condition. Average ratings for workload acceptability (1 = completely unacceptable, 5 = completely acceptable) were above the neutral point (i.e., 3) for all three conditions.

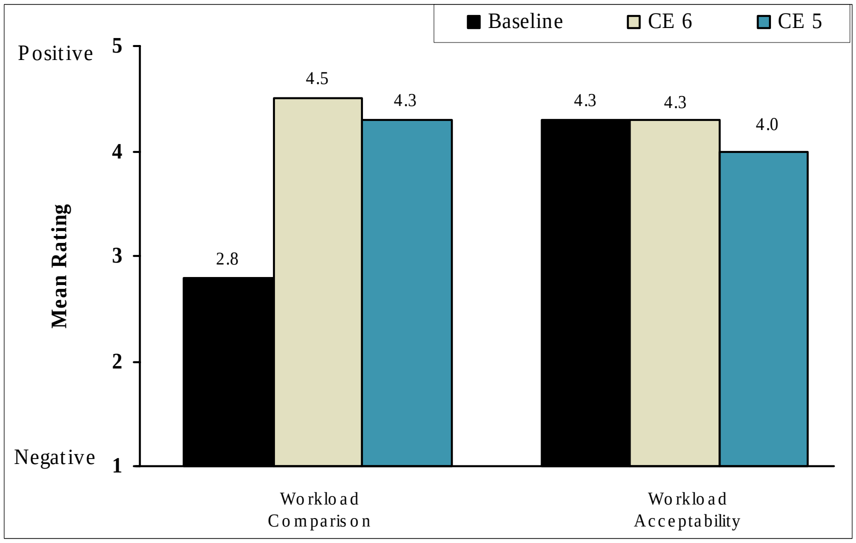

Figure 33. En Route Controller Workload Comparison and Acceptability Ratings According to Condition.

After each run, controllers rated the mental demand, effort, and frustration they perceived during that particular trial (1 = Low, 5 = High). As shown in Figure 34, average ratings for each of the three measures are highest in the Baseline condition, with average ratings for the CE 6 and CE 5 conditions at or below the neutral point.

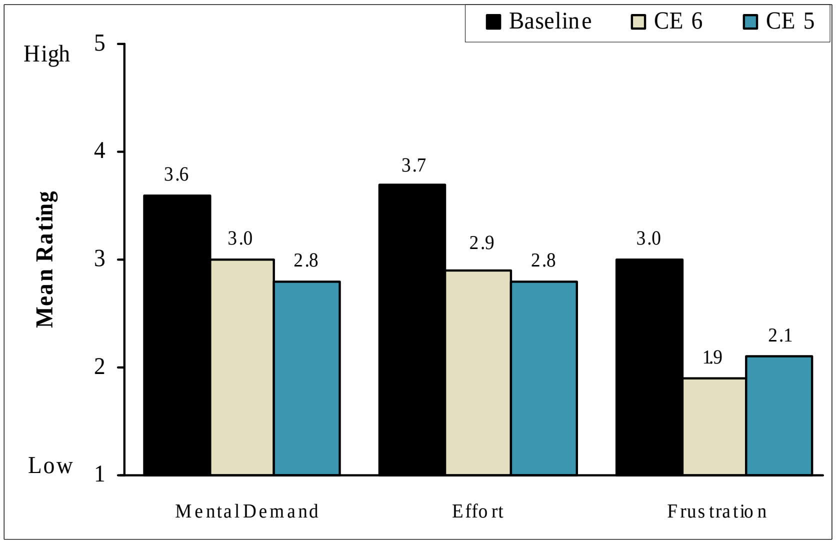

Figure 34. Controller Rating of Mental Demand, Effort, and Frustration According to Condition.

3.4.5 En Route Controller Situation Awareness

For each run, controllers rated how difficult it was for them to maintain situation awareness (1 = difficult, 5 = easy). The average ratings for the three conditions were above the neutral point and did not vary greatly between conditions (Figure 35). The highest ratings were for the CE 6 condition and the lowest for the CE 5 condition.

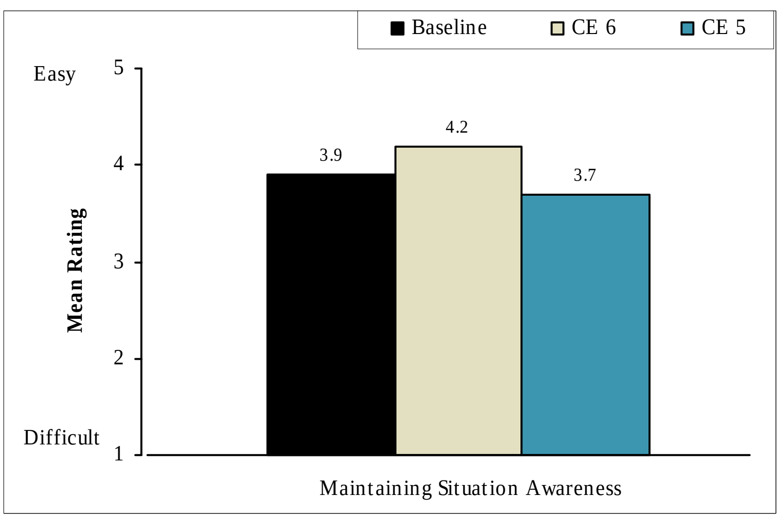

Figure 35. Controller Rating of Maintaining Situation Awareness According to Condition.

3.5 Cockpit Situation Display

3.5.1 Usability and Usefulness

Pilots evaluated the usability and usefulness of 28 CSD features by rating each one in the post-simulation questionnaire. Table 35 shows the average ratings and standard deviations. Only one feature, the approach spacing tool functions on the control strip, received an average rating of less than 3, the neutral point, for usability. The average suitability ratings for all CSD features were above 3.

Table 35. Pilot Usability and Usefulness Ratings of CSD Features.

	Usability (1=very difficult to use, 5=very easy to use)		Usefulness (1=unnecessary, 5=vital)		
Features	**Mean**	**Std Dev.**	**Mean**	**Std Dev.**	**N**
Aircraft Information					
Color coding of traffic w/ respect to altitude	4.4	0.74	4.1	1.13	8
Altitude tail tag (absolute)	4.4	0.98	4.6	0.79	7
Altitude tail tag (relative)	4.0	1.10	3.7	1.21	6
Altitude trend arrow	3.9	0.99	3.9	0.99	8
Data tags	4.3	0.71	4.3	0.71	8
Route Information					
Waypoint symbol	3.6	0.74	4.0	0.76	8
Waypoint color coding	3.4	0.92	3.5	0.76	8
Waypoint naming convention	3.5	0.93	3.6	0.52	8
Current waypoint information (name, time until reaching, and current distance to the waypoint) in top-right corner	3.1	1.68	3.4	1.41	7
4D flight plans	3.6	0.55	3.8	0.75	5
Flight path predictors (static mode)	3.9	0.64	3.6	0.92	8
Flight path predictors (pulse mode)	3.9	0.83	4.1	0.83	8
Flight path predictors (length)	3.5	0.93	4.0	0.93	8
Primary Alerting System					
Alert level color coding	4.0	1.07	4.3	0.89	8
Alert level auditory warnings	3.8	1.67	4.1	0.99	7
Alert level text warning	4.3	0.76	3.9	0.83	8
Route Analysis Tool					
Alternative flight path symbology	3.5	0.93	4.1	0.64	8
Waypoint table (bottom-right corner)	3.5	0.53	3.8	0.71	8

	Usability (1=very difficult to use, 5=very easy to use)		Usefulness (1=unnecessary, 5=vital)		
Features	**Mean**	**Std Dev.**	**Mean**	**Std Dev.**	**N**
Null point (white circle) and yellow cancellation line	3.1	0.83	3.1	0.64	8
RAT alerting system	3.6	0.52	3.6	0.74	8
RAT functions on control strip	3.1	0.90	3.7	0.95	7
RAT click and drag feature	3.3	0.89	3.6	1.06	8
Terminal Approach Spacing Tool					
Relative position of ownship to spacing box	4.1	0.83	4.1	0.64	8
Color coding of spacing box	4.4	0.74	4.3	0.71	8
Commanded speed in top-left corner	4.0	0.58	4.1	0.69	7
Approach spacing tool functions on control strip	2.6	1.06	3.5	1.20	8
User Settings					
Display setting (full mode [compass rose] and expanded mode [compass arc])	3.8	0.84	3.6	0.89	5
Range setting	3.5	0.76	3.5	0.93	8

Several questions on the post-simulation questionnaire dealt with specific usability and suitability issues (Table 36). Six pilots reported that the traffic display and primary alerting system effectively helped in maintaining safe flight operations (1 = very ineffective, 5 = very effective). The two other pilots had a neutral opinion in this regard. The average rating was 3.9 (*SD* = 0.64). Pilots rated whether the primary alerting system gave them enough time to resolve conflicts. Pilot ratings all have a value of at least 3, with an average rating of 3.8 (*SD* = 0.71). When asked how easy it was to distinguish the three levels of conflict alerting (1 = very difficult, 5 = very easy), pilots reported that they were easy to distinguish (*M* = 4.0, *SD* = 1.07).

One post-questionnaire query asked pilots to indicate whether the RAT allowed them to effectively construct alternative trajectories during the en route phase (1 = not at all, 5 = very much So). An average rating of 2.8 (*SD* = 1.04) suggests that pilots may not have been completely satisfied with the effectiveness of the RAT. Pilots also rated the process of implementing a flight path change with the RAT (1 = very Inefficient, 5 = very efficient). Four of the pilots gave a neutral rating of 3; however, opinions varied among the participants (*M* = 3, *SD* = 1.20).

One objective of data link communication is to reduce communication workload, so pilots were asked to rate the efficacy of the data link (1 = not at all Helpful, 5 = very Helpful). Responses were generally positive, with six out of seven pilots rating it at or above 3. The average rating was 3.7 (*SD* = 1.11).

Table 36. Pilot Ratings for Issues Related to CSD Use.

Question (Scale: 1=Negative, 5=Positive)	Mean	Std Dev.	N
Rate the effectiveness of the traffic display and primary alerting system for safe flight operations.	3.9	0.64	8
Did the primary alerting system provide enough time to strategically resolve conflicts?	3.8	0.71	8
How easy was it to distinguish the three levels of conflict alerting?	4.0	1.07	8
Did the Route Analysis Tool (RAT) allow you to effectively construct alternative trajectories during the en route phase?	2.8	1.04	8
Was the process of implementing a flight path change efficient (i.e., constructing a route, submitting it for approval, loading it into the FMS, and executing change)?	3.0	1.20	8
How helpful was the data link at reducing communication workload during the en route phase?	3.7	1.11	7

In response to the post-simulation questionnaire, pilots suggested additional information should be provided on the CSD. Suggestions were made for each of the three phases of flight (Table 37). The most common suggestion was to include the time of arrival of other aircraft in the data blocks.

Table 37. Additional CSD Information Requested by Pilots.

Phase of Flight	Information	n
En route	Time of arrival (e.g., PTA/ETA/STA) in the data block of other aircraft.	6
	Vertical speed of other aircraft.	2
	5- or 3-mile circles around selected aircraft.	1
Transition	Vertical speed on data blocks (relative mode) for own ship and intruder aircraft.	1
	STA and RTA of other aircraft.	1
Terminal	Distance to other aircraft.	1
	Runway position or actual outline.	1

Answers to the post-simulation questionnaire also indicate that pilots found some of the CSD information not to be useful or added clutter to the display. Table 38 lists these comments by phase of flight.

Table 38. Information Pilots Considered Not Useful or Added Clutter.

Phase of Flight	Information	n
En route	Aircraft at different altitudes. Only same altitude aircraft need be displayed.	1
	Countdown of seconds to next waypoints.	1
Transition	Route and waypoints of other aircraft selected when evaluating downstream conflicts.	1
	Other aircraft well above or below, which in no way affect your flight plan.	2
Terminal		

Comments concerning issues and problems with the flight deck tools were collected throughout the simulation runs and debriefings. Table 39 lists several of these comments. The number of pilots who contributed to each remark was not recorded.

Table 39. Pilot Comments Concerning Flight Deck Tools.

Phase of Flight	Comment	n
En Route	STA/RTA data link message contains the scheduled time to the meter fix; however, it does not provide the name of the meter fix.	-
	Both the STA and RTA were identified as “STA” in the data link message window.	-
Transition	“Free flight” in the data link message window should change to “ATC Controlled” upon entering terminal airspace.	-
Terminal	Pilots were uncomfortable with the slow rate at which aircraft moved to occupy the spacing box.	-
	The spacing symbology visually obstructed proximal traffic.	-

3.6 En Route Controller Workstation

3.6.1 Usability and Usefulness

In the post-simulation questionnaire, controllers rated several features of the controller workstation on usability and usefulness. Table 40 shows average controller ratings for the usability (1 = very difficult to use, 5 = very easy to use) of display features and tools and for the usefulness (1 = unnecessary, 5 = vital) of these items. All features were rated positively for usability except the conflict list and the route modification tool, which received neutral ratings. The only feature that received a negative rating for usefulness was the conflict list (*M* = 2.8).

Table 40. En Route Controller Rating of Usability and Usefulness of Displays and Tools.

	Usability (1=very difficult to use, 5=very easy to use)		Usefulness (1=unnecessary, 5=vital)		
Feature	**Mean**	**Std. Dev.**	**Mean**	**Std. Dev.**	**N**
Timelines	4.8	0.50	4.7	0.58	3
Shortcut window	4.2	0.96	4.5	0.58	4
Speed information in data block	4.8	0.50	4.5	0.58	4
Speed advisories	4.2	0.96	4.5	0.58	4
Conflict list	3.0	1.73	2.8	1.50	3
Trajectory preview	4.5	1.00	3.5	0.58	4
Coordination interface	3.8	0.96	3.8	0.96	4
Route modification tool	3.0	1.63	4.5	0.58	4
Keyboard functions	3.7	1.15	3.8	0.96	3
Mouse operations	3.8	1.26	4.3	0.50	4
Tool bar	4.0	1.41	3.8	1.26	4
Color coding of information	4.8	0.50	4.5	0.58	4

In addition, controllers indicated that the amount of clutter caused by the information presented on the display was somewhat unacceptable, *M* = 2.8 (1 = unacceptable clutter, 5 = not a problem). Controllers also indicated that they were able to distinguish autonomous from managed aircraft (*M* = 4.8), and that the interface for interacting with a mix of managed and autonomous aircraft was adequate (*M* = 4.3). Results are shown in Table 41.

Table 41. En Route Controller Ratings for Controller Workstation.

Question (Scale: 1=Negative, 5=Positive)	**Mean**	**Std Dev.**	**N**
Did the information presented on the display during CE 5 and CE 6 operations create clutter?	2.8	1.26	4
How easy was it to identify autonomous aircraft?	4.8	0.50	4
How adequate was the interface for working a mix of autonomous and managed aircraft?	4.3	0.96	4

3.7 Simulation Fidelity

Pilots found the simulation to be realistic in terms of communications with controllers. Specifically, seven out of eight pilots rated the realism of controller communication with the maximum score of five (1 = unrealistic, 5 = realistic). The other pilot rated it neutral, 3. Pilots rated the level of traffic complexity in the scenarios as comparable to that of normal current-day operations (1 = much lower, 5 = much higher). Results are presented in Table 42. Furthermore, to improve the realism of the simulation, pilots provided the suggestions listed in Table 43.

Table 42. Pilot Ratings of Simulation Fidelity.

Question (Scale: 1=Negative, 5=Positive)	Rating	Std Dev.	N
How realistic were communications with the controllers?	4.8	0.71	8
Rate the traffic complexity of this scenario (en route airspace).	2.8	0.69	8

Table 43. Pilot Comments for Improving Simulation Realism.

Comment	n
Use a full-motion simulator.	1
Use different voices for each pseudo-pilot flight.	1
Increase the number of potential conflicts in the terminal airspace.	1
Improve the quality of transmissions from ATC.	1
Have more runway changes occur.	1

Table 44 shows en route controller ratings of aspects related to simulation fidelity. Controllers rated traffic complexity as compared to normal current-day operations (1 = less complex, 5 = more complex). Average ratings for traffic complexity in all conditions were comparable to normal current-day operations. Controllers also rated traffic load (1 = light, 5 = heavy). Average ratings indicate that traffic load was perceived to be moderate. Finally, controllers rated the realism of the problem after each run (1 = unrealistic, 5 = realistic). Average ratings were slightly above the neutral point.

Table 44. En Controller Ratings of Simulation Fidelity.

Question (Scale: 1=Negative, 5=Positive)	Rating	Std Dev.	N
Rate traffic complexity compared to current operations.	2.5	0.87	4
Rate the traffic load.	3.2	0.81	4
Rate the realism of the problem.	3.5	0.97	4

When asked to provide recommendations for improving the fidelity of the simulation, controllers made the comments listed in Table 45.

Table 45. Controller Comments for Improving Simulation Fidelity.

Comment	n
Have pseudo-pilots maintain assigned speeds.	1
Provide more accurate delay times.	1
Speeds on FMS transition should be modified to slow aircraft to 210 knots at about 5 miles outside TRACON boundary.	1
Have pilots refuse speeds at the back of the envelope.	1
Have pilots be more responsive to clearances and communication changes.	1
Have autonomous aircraft at the end of the problem.	1
PCPlane aircraft slowed at a different rate than the pseudo-pilot aircraft (CE 11).	1

3.8 PARTICIPANT TRAINING

Pilots indicated that the training they received at the beginning of the study could be improved. In the post-simulation questionnaire, pilots rated the adequacy of six aspects of the training, as listed in Table 46. Average ratings ranged from 2.6 through 3.4 (1 = not adequate, 5 = adequate). One pilot suggested that the training be more task-oriented, while another felt that PCPlane procedures should first be demonstrated on a workstation.

Table 46. Pilot Ratings for Adequacy of Training.

Training Element (Scale: 1=Negative, 5=Positive)	Rating	Std Dev.	N
Separation responsibility and rules-of-the-road	2.6	0.52	8
Display symbology	2.8	0.89	8
Procedures for interacting with PC Plane	2.8	1.28	8
Phraseology	3.1	0.64	8
Concept familiarization	3.3	0.89	8
ATC procedures	3.4	1.19	8

En route controllers rated the overall adequacy of the training they received (1 = completely inadequate, 5 = completely adequate). The mean rating, 4.1 (*SD* = 0.50), was above the neutral point.

4. DISCUSSION

4.1 SYSTEM PERFORMANCE – EN ROUTE AIRSPACE

4.1.1 Efficiency

Several metrics were used to measure aspects of system efficiency in the en route environment. System-level metrics focused on the initial descent phase of flight, encompassing en route operations from approximately 200 nm outside the TRACON boundary to the meter fix. Metrics include absolute arrival delivery error, en route altitude, altitude deviation (at the meter fix), distance traveled, and flight time.

Absolute Arrival Delivery Accuracy

Absolute arrival delivery error was a measure used to determine how accurately controllers and pilots of autonomous aircraft could maintain the arrival schedule. The assumption is that system efficiency would benefit from an accurate feed to the TRACON. Absolute arrival delivery error varied significantly across conditions. En route controllers were able to provide a steady and efficient feed to the TRACON controller in the CE 6 and CE 5 conditions, as indicated by a small absolute arrival delivery error (17.5 and 17.1 seconds, respectively). However, in the Baseline condition, the feed to the TRACON was less representative of the actual arrival schedule, with an absolute arrival delivery error of 59.9 seconds.

One explanation for an increase in absolute delivery error in the Baseline condition is the smaller number of groundside tools in that condition. Controllers operating in the Baseline condition only had the TMA sequence list and delay information available to them. They did not have access to the TMA timelines, DA, CPDLC, or conflict detection and trial planning tools. The TMA timeline, which displays aircraft STA and ETA on side-by-side timelines, likely helped directly in management of the arrival schedule. CPDLC, conflict detection, and trial planning tools may have also supported the arrival schedule by enabling controllers to strategically manage traffic in their sector. Specifically, the trial planning tool allowed the controller to assess the impact a heading or speed change would have on the arrival schedule before issuing the clearance. Not surprisingly, when asked to rate their difficulty in managing the arrival schedule, en route controllers consistently rated schedule maintenance as less difficult under CE 6 and CE5 conditions than the Baseline condition.

Another factor likely contributing to the increase in absolute delivery error in the Baseline condition is that the controllers generally did not use the swap function on the PGUI toolbar. When two aircraft are physically swapped in the arrival sequence, typically by issuing heading or speed clearances, the controller must also enter the two flights IDs into the swap function so that the system can update the STAs. Otherwise, the simulation software will record the flights as out of sequence and compute a delivery error based on the original STA. A disproportionate number of resequenced flights were not entered with the swap function during Baseline runs, which inflated the absolute delivery error. At the time of publication, it has not been determined why controllers tended not to use the swap function in the Baseline condition. (See Section 4.1.3 for a discussion of learning effects.)

Altitude Deviation

Flights deviating from their assigned crossing altitudes have the potential, depending on the magnitude of the deviations, to reduce system efficiency and create extra workload for controllers. The magnitude of the deviations reflects, in part, the effectiveness of the arrival procedures. During the en route descent phase, the CE conditions utilized the precision descent procedure, while the Baseline condition used standard descent procedures.

There were no significant differences between conditions for average altitude deviations for flights that were outside the meter fix crossing buffer (±299 feet). However, the proportion of flights that were off-altitude was significantly larger in the Baseline condition than in the CE 6 and CE 5 conditions. It appears that aircraft were as good, if not better, at meeting the meter fix crossing restriction when flying the FMS-generated precision descent (CE 6 and CE 5 conditions) as when relying on controller clearances (Baseline condition).

En Route Altitude

Because a flight is generally more fuel efficient at higher altitudes, flight efficiency was investigated by looking at en route altitude. The average en route altitude at 160 nm from the meter fix did not vary significantly by condition. This was likely due to the fact that most flights were still at their cruise altitude. A significant difference appeared between conditions at 60nm from the meter fix, which constitutes the initial descent phase. The lowest average altitude was found in the Baseline condition.

The difference in average altitude is likely a result of the different descent procedures. All aircraft in the CE 6 and CE 5 conditions utilized the precision descent procedure that permitted pilots to fly the aircraft's FMS descent profile. Generally, this allowed aircraft to maintain cruise altitude longer and start their descent later in the flight than aircraft in the Baseline condition, which started their descent earlier due to ATC directives.

Distance Traveled

Distance traveled to the meter fix was recorded from when a flight first reached the 160-nm arc. Differences in distance traveled were found to be significant. Results showed that flights in the Baseline condition flew, on average, 7.4 and 9.0 more miles than in the CE 6 and CE 5 conditions, respectively. The analysis of PCPlane flights showed a similar trend (but not a significant one), with slightly larger differences between conditions. The shorter distances flown by PCPlane flights, particularly in the CE 5 condition (164.5 nm), may have only slightly contributed to the overall trend because PCPlane flights accounted for a relatively small proportion of the total arrivals.[2]

As mentioned previously, differences in groundside tools may have contributed to the greater difficulty controllers had in meeting the arrival schedule in the Baseline condition. This may also be true for distance traveled. The speed advisory tool was not available to controllers in the Baseline condition, which may have led to more vectoring, as opposed to speed changes, when aircraft needed to absorb delay. Similarly, the route-planning tool was not available, so controllers could not preview the amount of delay a vector would absorb. Controllers had to learn by trial and error in the Baseline condition, which may have produced larger than necessary vectors. This seems to be supported by the higher variance in arrival delivery error in the Baseline condition. Controllers were not as proficient in adjusting a flight plan to meet the STA. (See Section 4.1.3 for a discussion of learning effects.)

[2] Approximately 18 percent of DFW arrivals were flights flown via PCPlane workstations.

Flight Time

Flight time to the meter fix was recorded from when a flight first reached the 160-nm arc. Results showed that flight time varied significantly across conditions, with average flight times in the CE 6 (1.1 minutes) and CE 5 (1.6 minutes) conditions less than in the Baseline condition. Not surprisingly, these results are consistent with the distance-traveled results. The additional miles traveled in the Baseline condition likely impacted flight time correspondingly. Another explanation for the increase in flight time in the Baseline condition is the descent procedures used. The Baseline condition used standard descent procedures that resulted in aircraft being descended earlier and stepped down throughout the descent. In both cases, these procedural changes slowed an aircraft's initial descent to the meter fix as compared to the precision descent procedure. First, the early descent took aircraft out of cruise speed sooner and, second, the series of step-downs required aircraft to slow in order to level off each time.

4.1.2 Safety

Two system-level metrics used during the simulation, separation violations and arrival spacing, provided a measure of overall safety. Separation violations were counted only if they occurred within or between sectors staffed by participant controllers. Arrival spacing was analyzed to determine the extent to which controllers and pilots deviated from the 58-second crossing minimum (which equated, approximately, to the minimum separation requirement). Flights with a spacing interval less than 58 seconds could be in violation of the separation minimum (if not separated by altitude) and therefore posed a safety problem.

Separation Violations

The results suggest that CE 5 and CE 6 operations did not decrease overall safety, as indicated by fewer separation violations occurring in the CE conditions than in the Baseline. Overall, only one violation occurred near the meter fix where the heaviest congestion existed, suggesting that en route controllers and pilots of autonomous aircraft were able to adequately handle the traffic load. Two out of nine separation violations involved a PCPlane flight, with both occurring in the Baseline condition. Because neither violation occurred in a CE condition, the results further support the viability of the DAG-TM operational concepts.

Furthermore, safety in the high altitude sectors may have been increased, possibly due to the added functionalities of the ground- and airside tools. For example, there were three violations in the Baseline condition between one aircraft in a descent phase and another in a climb phase. On the basis of controller comments, it seemed that the ability to differentiate between departures and arrivals was enhanced in the CE conditions by the color coding of aircraft targets (yellow for arrivals and green for departures). Perhaps, this coding allowed controllers to identify departures and arrivals on crossing paths more quickly than without the coding, which afforded them more time to take corrective actions.

These findings should be viewed as conservative measures, and a lower number of violations would have likely occurred if the controllers were given equivalent support to that which is available in the field today. For example, in the Baseline condition, controllers were operating without the aid of a conflict alert tool and a D-side controller.

Arrival Spacing

When taking into account only the flights that deviated from the 58-second separation minimum at the meter fix, results showed that the proportion of flights varied significantly between conditions, with the greatest proportion occurring in the Baseline condition. These findings suggest that controllers had less difficulty in maintaining the minimum spacing requirement during the CE 6 and CE 5 runs than during Baseline runs. This finding indicates that safety was enhanced through the use of DAG-TM tools and procedures. Both separation violation and arrival spacing data upheld the DAG-TM principle that gains in efficiency and flexibility should not negatively impact safety.

4.1.3 Learning Effect

Several of the system performance metrics may have shown a learning effect for the Baseline condition only. These dependent measures, which include arrival delivery error, distance traveled, and flight time, reflect the team's ability to adhere to the arrival schedule. All but one controller were from an ATC facility that does not perform arrival metering, so the task of meeting the arrival schedule was a new challenge for the controller team. The CE 6 and CE 5 conditions provide several specialized tools for managing the arrival schedule. In particular, controllers had speed advisories, a route-planning tool, and the TMA timelines available to them. These tools allowed controllers to trial plan before issuing a flight plan change.

In the Baseline condition, controllers had no way to preview the impact a flight plan modification would have on the arrival schedule. They had to learn how to absorb delay without immediate feedback. This meant learning, through trial-and-error, how to judge the magnitude of a vector or speed change in terms of arrival delay. Baseline results showed that performance generally improved throughout the data collection period, as controllers became more proficient at managing the arrival schedule. In contrast, results of the CE 6 and CE 5 conditions did not show a similar trend. It is possible that a learning effect did not occur because controllers had previous experience with the DAG-TM DSTs.

4.2 SYSTEM PERFORMANCE – TERMINAL AIRSPACE

Self-spacing within the terminal airspace was reasonably successful. There were no separation violations involving PCPlane aircraft that were self-spacing. The closure rate for these aircraft was consistent for both CE 6 and CE 6 flights even though CE 5 flights averaged nearly a minute longer in total spacing time.

Initially, it appears that the procedures and tools enabled the controller and pilots to provide safe and reliable self-spacing throughout the operation. However, the conservative technique used by the TRACON controller may have been primarily responsible for these results. As mentioned, the controller developed his own style of implementing the self-spacing task that was different from the procedures suggested during training. Although he was instructed to assign 90-second spacing intervals, he regularly assigned a spacing interval that was slightly greater than what the aircraft was actually doing. For this reason, the average temporal spacing change was positive.

Furthermore, self-spacing execution, on instructions from the TRACON controller, was inconsistently applied. There were times when spacing was begun at a reasonable distance from the FAF; but at other times, it bordered on being so late in the approach as to be irrelevant. There are two possible explanations for the controller's behavior. First, he was communicating spacing instructions to aircraft when workload permitted; and second, he only gave a spacing clearance when the aircraft-to-aircraft relationship was such that the overall flow would not be disrupted. Although the issue of workload remains a fundamental concern, implementing self-spacing only in limited situations will not achieve the benefits the concept was designed to provide.

An interesting finding was the difference in total time spacing between the CE 6 and CE 5 conditions. In theory, there should have been no difference (in mean total time spacing) because self-spacing operations within terminal airspace were the same for the two conditions. Furthermore, arrival spacing and altitude deviations at the meter fix were nearly identical for the two CE conditions, suggesting that the TRACON controller was getting similar feeds from the low-altitude controller. One explanation may be that, because the controller was learning and adjusting throughout the simulation, performance also changed (and the last run was the CE 5 condition).

At a minimum, data on inter-arrival spacing times for non-self-spacing aircraft should be collected in future simulations to provide a baseline measure in which to evaluate the CE 11 concept against.

4.3 WORKLOAD AND SITUATION AWARENESS

4.3.1 Pilot

The descriptive analyses of the subjective judgments of workload made by pilots after each run and at the end of the simulation study suggest that perceived workload may not have varied between conditions, and that it was perceived to be comparable to, or slightly lower than, the level experienced during normal current-day operations. A few comments made during debriefings and captured through the post-simulation questionnaire suggest that pilot workload may have increased when both conflicting aircraft pilots were trying to resolve the situation. Another pilot mentioned that workload increased as he approached the meter fix.

Similarly, pilots said that their level of SA was comparable to normal current-day operations or even slightly better. SA estimates did not seem to vary greatly between modes of operation.

One rationale for why subjective judgments of workload and SA seem not to have varied between conditions may be that the level of complexity or difficulty of the scenarios was not high enough for the pilots to need the additional help provided by the CE 6 and CE 5 tools and procedures. Pilots’ ratings of complexity, workload, and SA suggest that complexity during the simulation runs was comparable or even slightly lower than in normal current-day operations. Differences between conditions may not have emerged because pilots had not used the tools often enough during each run to have an impact on their overall assessment of workload and SA. While the absence of workload increase or SA deterioration is a positive outcome, future investigations would benefit from examining these factors at different moments during each run and by employing objective assessment methods.

4.3.2 En Route Controller

Post-run questionnaire responses suggest that en route controllers maintained an acceptable level of situation awareness during the simulation and that it did not vary across conditions. The same controllers reported that workload in the Baseline condition was comparable to what they usually experience in normal current-day operations, and that it was less in the CE 6 and CE 5 conditions. The analysis of the ATWIT ratings revealed a similar pattern of response, although the rating differences between the three conditions were only statistically significant at the Amarillo position. The controllers also found that the level of workload in all three conditions was acceptable.

The subjective workload assessments suggest that the observed improvement in system performance in the CE 6 and CE 5 conditions did not diminish en route controller situation awareness and may have reduced controller workload. The presence of ground decision support tools, such as TMA timelines, DA, CPDLC, and conflict detection and trial planning, probably explains why controllers found it easier to accomplish their tasks in the CE 6 and CE 5 conditions. These tools may have, for example, helped controllers to maintain the arrival schedule, facilitated strategic planning, reduced the need for vectoring, and helped determine the amount of delay a flight path change would absorb.

4.4 CONCEPT ELEMENT ISSUES

4.4.1 Cancellation of Autonomous Control

The option to cancel autonomous control in the CE 5 condition was intended to ensure safe and efficient operations. Either ATC or the flight deck could initiate the cancellation of autonomous control due to traffic, equipment or scheduling problems, or other extenuating circumstances. ATC cancellation of autonomous control was generally considered a negative event from the flight deck perspective, although cancellation of autonomous control does not always have to be. For example, ATC could cancel autonomous control to give an aircraft a direct route to the meter fix. This raises two questions: Should ATC be allowed to cancel autonomous control if a direct route to the meter fix is available? Should the pilot be able to refuse, for example, if an early arrival creates a gate availability problem?

The cancellation of autonomous control was a relatively rare event, occurring only twice in 28 flights. However, when it occurred, the question arose as to whether autonomous control should be reinstated or whether the aircraft should remain under the controller's control. The one time autonomous control was reinstated occurred during the initial descent phase. The pilot of an autonomous aircraft requested that the controller take control because of a pending traffic conflict. The controller assumed control and resolved the conflict. Afterwards, the pilot requested, and was granted, autonomous control. However, the aircraft's proximity to the meter fix negated the benefits of free maneuvering. In the end, the actions only increased pilot and controller workload.

It was suggested that the resumption of autonomous control should not occur if the aircraft does not have the opportunity to benefit from free maneuvering. Two points along the flight path were suggested as possible cut-off points in which autonomous control should not be reinstated: top-of-descent and the TRACON merge point (e.g., UKW in the current study).

4.4.2 Assigned Arrival Time

All arriving aircraft were constrained by an assigned arrival time that was conveyed to the pilot as STA and RTA clearances. Aircraft whether free maneuvering or negotiating trajectory changes with ATC, were instructed to meet these arrival times. The STA was issued to the aircraft before it reached the TMA freeze horizon, approximately 200-300 miles from the destination airport. The RTA was issued once the aircraft reached the freeze horizon. One problem was that both the STA and RTA were identified as "STA" in the data-linked message. This designation made it difficult for pilots to distinguish between the two time clearances. It was important for pilots to differentiate the two clearances because the STA permitted them to adjust their proposed time of arrival (PTA) and the RTA did not. Updating the PTA after the RTA was received caused large discrepancies between flight deck and ATC ETAs.

Procedures Concerning the Assigned Arrival Time

Several procedural issues arose concerning the assigned arrival time. In the CE 6 condition, ATC occasionally sent speed clearances that took equipped aircraft off their STA/RTA. Pilots were not clear whether they should still try to make the STA/RTA, which could be accomplished by modifying their flight paths. During the debrief, it was decided that in these situations, the STA/RTA should no longer be in effect and the controller should inform the pilot that the STA/RTA has been invalidated.

When a pilot made a flight path change to avoid a LOS, the maneuver often took the aircraft off its RTA. The controller would then swap the aircraft's arrival slot to provide conflict-free metering. The question is whether a new RTA should be assigned after such a route modification, and if so, how? Should the pilot request a revised RTA from the controller or will the controller automatically send it once a new time is assigned?

Also, in the CE 6 condition, a question arose about whether equipped aircraft are required to fly the assigned speed once cleared for the precision descent. It appeared that the CTAS descent advisor did not always correspond to the FMS descent profile, so the speed assigned was not optimal for meeting the PTA. During the debrief, it was suggested that pilots be allowed to make speed changes without ATC concurrence in order to meet their PTA. However, it was not established whether limits were necessary to govern how much the speed could vary from the descent clearance speed.

4.4.3 Rules-of-the-Road

Pilot

Pilot ratings regarding ease-of-use of the rules-of-the-road were slightly favorable. In terms of safety, pilots rated the rules-of-the-road as resulting in a level of safety that was slightly higher than the level of safety provided by current-day procedures. Pilot ratings indicated that, when they experienced a Level 3 alert, they were relatively certain as to who was responsible for maneuvering, especially in the CE 6 condition. Because pilots were always under ATC control in the CE 6 condition, it is not unexpected that they were certain as to who was responsible for maneuvering. Pilot ratings were lower for the CE 5 condition than for the CE 6 condition. Less certainty is to be expected because either aircraft or ATC could be responsible for resolution in the CE 5 condition.

Although the pilots understood the basic rules, they had trouble applying the rules-of-the-road to some specific conflict geometries. Some of the confusion on the pilots' part may be attributed to lack of training or experience with the rules. However, the following two difficulties represented reoccurring problems – trouble determining whether traffic aircraft were within 20° of own ship and trouble determining who had the right-of-way in certain overtake situations.

Debrief discussions concerning rules-of-the-road and resolution responsibility resulted in the group consensus that information should be provided to indicate which aircraft is responsible for resolving the conflict. This information would preclude pilots and controllers from having to make this decision, thereby reducing cognitive workload. Providing this information may reduce the ambiguity that currently exists in the system and result in more timely and efficient maneuvers.

En Route Controller

Answers to the post-run and post-simulation questionnaires show that en route controllers gave a slightly favorable rating to the adequacy of the rules-of-the-road and the associated procedures and phraseology. Two controllers mentioned in the post-simulation questionnaire that more training focusing on the rules-of-the-road would be necessary to establish an efficient distributed control environment. A possible lack of familiarity with the rules may not have impacted conflict resolution events much because the number of times the rules were applicable (only in the CE 5 condition) was limited and autonomous flights and controllers were permitted to resolve conflicts even if they were not responsible.

For example, one controller described a situation where, by the time an autonomous aircraft responsible for resolving a conflict moved, the controller had already sent a resolution clearance to the managed aircraft. Although controllers may have found that the rules were efficient in determining resolution responsibility, this example illustrates a potential problem. The rules-of-the-road established which user was responsible for resolving a conflict without imposing a specific resolution strategy. These rules offered users the flexibility to adopt the strategy they preferred. This may have caused some problems in situations where the aircraft not responsible for resolving the conflict acted, while the other aircraft was implementing its own resolution strategy. Unanticipated movements could increase the complexity of the situation, or even raise the likelihood of a separation violation.

Further investigations should consider whether rules-of-the-road should also establish specific conflict resolution strategies, as *Federal Aviation Regulations* (FAR) Part 91 (FAA, 1997) right-of-way-rules do. This section and Section 4.4.2 emphasize the need to design rules or a system that will clearly establish conflict resolution responsibility; but these assignments will also need to correlate with user needs (e.g., offer resolution strategies) and optimize system efficiency.

4.4.4 Conflict Resolution

Both pilots and ATSPs indicated that they had several concerns about resolving traffic conflicts in the CE 5 condition. The primary concerns were determining who was responsible for resolving the conflict and knowing that the individual responsible had recognized the conflict and started working on a resolution. Pilots also indicated that help with resolving traffic conflicts would be beneficial, particularly with Level 3 alerts (which present the shortest time to LOS and require immediate action).

Procedures for Conflict Resolution

Events occurred during the CE 5 runs that led to discussions about changes to the resolution procedures. There are cases where alternative procedures may be preferable to those prescribed in the rules-of-the-road, which state that the individual responsible must maneuver. Pilots mentioned that it could be beneficial to negotiate a resolution. For example, there might be a situation in which an aircraft that is required to maneuver, based on conflict geometry, could be on time while the aircraft with the right-of-way could be early. In this case, should the pilot of the on-time aircraft have the right to request that the other aircraft maneuver instead? Still, in other circumstances, should a pilot who is not responsible for maneuvering initiate a conflict resolution without consulting the other pilot or ATSP?

In addition, there were conflict situations that developed quickly or that pilots could not resolve in a timely manner. In one case, pilots declared an emergency and relinquished responsibility to ATSPs with 32 seconds to LOS. This event prompted a discussion as to what was a reasonable time frame in which pilots could pass responsibility for separation assurance to the ATSP. Obviously, passing a conflict off to the ATSP with half a minute to LOS is not acceptable. ATSPs voiced concerned that pilots, after failing to resolve a conflict, would hand the problem to them at the last minute, with one controller stating, “No one likes to have a hot sack of … grits dropped in their lap.” Although there was no consensus as to what the procedures should be, it was agreed that research is needed to develop procedures for such situations.

Information Requirements

Pilots could detect that an aircraft had initiated a maneuver only by monitoring the aircraft symbols on the CSD and looking for changes. If an alert was cycling on and off, pilots could not quickly determine if an alert disappeared because the other aircraft had maneuvered or the alert was in another off cycle. Uncertainty about whether the other pilot was actively working to resolve the problem was a major concern for pilots. This lack of feedback may be just as significant with conflicts between managed and autonomous aircraft, in which the controller is responsible for resolving the conflict. Controllers and pilots have different resolution levels and tolerances for acting on conflict situations and, up to now; controllers have more experience than their counterparts resolving conflicts. A problem may often look much more threatening to a pilot than to a controller. Pilots in the study tended to act promptly to resolve conflicts, while at times, controllers were apt to let the conflict unfold before resolving. If ATSP delays in resolving the conflict, the pilot may experience an increase in stress and workload (e.g., if the pilot starts to resolve the conflict on his/her own). For the CE 5 condition, both pilots and controllers desired confirmation that the individual responsible has detected the conflict and is aware that he or she is required to maneuver. In addition, they expressed a desire to know that the individual responsible is actually taking steps to resolve it.

Another concern for pilots was the situation when two aircraft were in conflict as they approached a common merge point. Specifically, in a case where own ship was responsible for maneuvering, the pilots could not make an informed decision as to whether they should maneuver in front of or behind the conflicting aircraft. Pilots requested that they be provided with information regarding the arrival status of the conflicting aircraft (i.e., whether the conflicting aircraft is early, late, or on time), regardless if managed or autonomous. Future studies could evaluate the benefits of providing this type of information to pilots and examine how the information should be presented.

Pilots also expressed concern about the difficulty they had in creating viable trial routes with the RAT. This was especially true with Level 3 alerts or when two conflicts needed to be resolved simultaneously. One suggestion was to provide automated resolution capability to the primary alerting tool, possibly offering pilots a few resolution alternatives (e.g., most fuel efficient or most expeditious). From a flight deck perspective, pilots felt that automated resolution would greatly facilitate the resolution process.

4.4.5 Terminal Approach Spacing

Procedures

Overall, pilots found the procedures for the approach spacing to be acceptable. Most pilots reported that the terminal approach spacing procedure was more efficient than current-day operations. Half of the pilots thought that the procedures were safer than current-day operations, and half were neutral.

However, some pilots expressed concerns about self-spacing spacing being time based, rather than distance based. They felt this increased the difficulty in maintaining an acceptable distance from the lead aircraft. There were other concerns about the ability of the auto-throttle to maintain the required speed and therefore the appropriate separation from the lead aircraft. Some pilots were uncomfortable with the slow rate at which aircraft moved to occupy the spacing box, especially if ATSP changed the spacing interval during final approach. In some cases, the lead aircraft would land before own ship ever achieved proper spacing.

One difficulty the TRACON controller had was determining when to initiate self-spacing. He felt pilots had a difficult time reacting in a timely way to the spacing assignment. For this reason he issued self-spacing intervals that reflected the aircraft's current spacing. In addition, the controller often waited until aircraft were well within terminal airspace before issuing the self-spacing clearance. At that point, in-trail spacing had typically become more stable and, perhaps then, the controller was more confident that the pilots could manage spacing themselves.

It is clear that the self-spacing procedures and their implementation need to be better defined for the TRACON controller. In future studies, researchers will need to formalize procedures and oversee the implementation of the task to ensure self-spacing is applied consistently.

Phraseology

The voice clearance used in the simulation to designate a lead aircraft confused some pilots. For example, a pilot whose aircraft was being assigned as lead would hear his or her callsign twice when not actually being called at all – once when ATSP assigned the aircraft as lead and once when the pilot of the trailing aircraft read the clearance back. This procedure was confusing for the pilot of the lead aircraft. It was suggested that the approach spacing clearance could be data-linked to pilots, which would eliminate the potential for confusion.

Information Requirements

Pilots indicated that the information they received was adequate accomplishing their tasks and that there was no unnecessary clutter on the display. However, some suggested that the data block of the lead aircraft should automatically display when the spacing tool engaged. This would enable the pilot to more easily locate and monitor the ground speed and altitude of the lead aircraft.

4.4.6 Flight Crew Coordination

During the study, the majority of participant pilots used the PC-based PCPlane, essentially playing the role of both captain and first officer. Only the two participant pilots in the ACFS were able to act as a coordinated team, with one pilot taking the role of captain and the other as first officer. This setup allowed ACFS pilots to distribute tasks between themselves and to collaboratively work out problems. As such, the comments presented here represent those observed during the runs in the ACFS only. We recognize that further study must be conducted on flight crew coordination within the DAG-TM environment.

Pilots in the ACFS did not receive training on how to delegate tasks. After a few runs, however, the pilots in the ACFS adapted to their individual roles and developed procedures to complete the required tasks. If a new situation developed, the pilots would distribute the tasks at that time.

For the most part, the pilots' efforts were well coordinated and executed effectively. However, one event illustrated how critical it was for each pilot to know what the other pilot was doing, so that they did not accidentally negate the other's efforts. On one occasion, the first officer was creating a trial route with the RAT while ATSP was sending up a new route. At that point, the first officer had not yet executed the trial plan on the RAT. Meanwhile, the captain loaded and executed the ATSP route on the FMS. The active flight plans changed on both pilots' CSDs, deleting the trial route on the first officer's CSD. When the new route appeared on the CSDs, it happened to be similar to the trial plan that the first officer had created. The pilots were not sure whether the newly loaded route was the ATSP route or the first officer's trial route.

4.5 ADVANCED DAG-TM TECHNOLOGIES

4.5.1 Cockpit Situation Display

Aircraft and Route Information

In general, the CSD was rated favorably. Several aircraft and route information features were positively rated in terms of usability and usefulness. For example, color coding of traffic with respect to altitude, altitude tail tags, data tags, altitude trend arrow, waypoint symbology and color coding, and flight path predictors all received ratings above neutral. A couple of exceptions were waypoint information and the 4D-flight plan. They were rated slightly negatively for usability, though they received positive ratings for usefulness.

One reason for the lower rating for the 4D-flight plan might have been the difficulty viewing altitude information when the flight path predictors were activated. An aircraft's altitude is indicated by its 4D-flight plan; color representing altitude relative to own ship and type of line, dashed or solid, representing a level or changing flight segment. However, when the flight path predictor is activated and extended, the predictor lines overlay the 4D-flight plan. Because the flight path predictors do not convey altitude information, altitude information about traffic aircraft is temporarily lost until the predictors are turned off.

Another reason for the less-than-favorable rating may be training related. Some pilots were observed questioning the 4D flight plan symbology during data collection runs. Specifically, they were unclear how altitude was being represented by the flight plan. Pilots specified training as an area for improvement (see Section 4.7.1), hence, the 4D-flight plan usability rating may have been more reflective of a shortfall in training rather than shortcomings in the tool itself.

Waypoint information included the time until reaching the current waypoint (in seconds). One reason waypoint information received a lower rating might have been the format used. Some pilots suggested that it would be more useful if it were displayed in Zulu time rather than seconds. However, at least one pilot used the "time to current waypoint" and descent rate for predicting the waypoint crossing altitude.

Primary Alerting System

Primary alerting system was judged as providing clear and timely alerting to potential traffic conflicts. Alert color-coding and auditory and text warnings were all rated favorably for usability and usefulness. Pilots also felt that the CSD, as a whole, helped in detecting potential conflicts prior to a system alert or ATC warning.

One reservation pilots had about the alerting system were its sensitivity. Pilots perceived the alerting logic as being over sensitive and at times inconsistent. Level 1 alerts would sometimes appear and disappear two or three times within a minute. Other times, the alert would first appear as a Level 3. Yet the frequency of these occurrences was relatively low with respect to the total number of alerts and, as mentioned, pilots were quite positive about the alerting system.

Route Analysis Tool

All RAT features were rated favorably for usability and usefulness. However, pilot responses were slightly negative and neutral, respectively, when asked to rate the effectiveness of the tool in constructing alternative flight paths and the efficiency in doing so. The following feedback from pilots about the functionality of the RAT may account for the relatively low ratings for effectiveness and efficiency.

One thing that complicated RAT use was that the RAT did not probe altitude and speed modifications, which had to be entered via the FMS, for potential conflicts's alerting logic. Pilots occasionally discovered that a new conflict was created after the modification was executed in the FMS. Pilots who initiated such changes through the FMS were occasionally required to modify the new route soon after because of a newly created conflict. If altitude or speed changes had been probed, pilots would have received a RAT alert prior to executing the new flight plan and may not have maneuvered into a new conflict situation.

Pilots also found it difficult to plan routes that adhered to the schedule because the ETA remained static during trial planning. Pilots expected that the ETA would reflect the modified route while the RAT was still in trial mode. However, the ETA updated to reflect a route modification only after the route was executed.

In addition, pilots indicated that the procedures for creating trial routes with the RAT were somewhat cumbersome. The task of creating waypoints and dragging waypoints until the conflict alerting symbology changed colors was a trial-and-error process. This process did not appear to cause problems in cases where there was a long lead-time to LOS or if pilots were initiating a route change for scheduling purposes, but pilots were sometimes under time pressure in cases where LOS was imminent. In these cases, it was challenging to create a single feasible route; and usually they did not have enough time to test alternative routes. In part, some of these problems may be attributed to the pilots' lack of experience with the tools, system, and strategies for resolving conflicts. If pilots had more experience in resolving specific conflict situations, they may have had an easier time creating trial plans.

Terminal Area Approach Spacing Tool

Pilots rated all but one of the features of the approach spacing tool favorably (e.g., color-coding and relative position of spacing box, and commanded speed). The one exception was the control strip functions, about which a few pilots also expressed concern about the number of steps and the time required to input spacing parameters. Specifically, they thought that if a change was required for one parameter, such as the spacing interval, they should not be required to reenter all the spacing information. Reentering information created extra workload for pilots operating in terminal airspace.

Pilots indicated that the spacing box symbology was easy to use; however, they were not clear on how to adjust speed to maintain the assigned interval when own ship was not spacing properly. Pilots felt the PDA/auto-throttle would prevent them from making manual speed changes, but the auto-throttle appeared to do an inadequate job of spacing. Several times the pilot engaged the spacing tool and properly spaced at the start of the approach. However, after a short time, aircraft would "drift" in front of or behind the spacing box even though the auto-throttle was engaged. Eventually aircraft would get back into the spacing box. Some pilots questioned why the approach spacing tool, coupled to the auto-throttle, could not maintain more consistent spacing.

One recommendation that may aid in self-spacing task is to replace the spacing box with a chevron-like target. This would require the pilot to position

User Settings

One difficulty pilots experienced were detecting traffic conflicts that were not within the limits of the current range setting. When more than one conflict existed, it was possible to miss the more distant conflict if the range setting was not great enough. As a result, pilots might not have been aware of a conflict that they were responsible for resolving until it reached a Level 3 alert. Zooming out to see a more distant conflict was not effective because the higher range setting produced an overly cluttered display. A suggestion made during the debrief was to incorporate a “scroll” feature that would allow the pilot to view a different area of the map display without having to zoom out.

ACFS

Both pilots in the ACFS were concerned about the amount of time they spent heads-down. The pilots indicated that the amount of heads-down time might be acceptable in the en route phase, but not in critical flight phases such as terminal approach. Also, pilots in the ACFS used touch pads and had great difficulty interfacing with the RAT due to the sensitivity of the input device. Performance was improved when the touch pads were replaced with a mouse and trackball. Even when using the mouse or trackball, pilots expressed concern about the usability of such sensitive input devices in actual flight conditions, especially during turbulence.

4.5.2 Controller Workstation

Overall, the four sector controllers rated the tools and interface as usable and useful. Some controllers actually mentioned that the tools made the task *too* easy. The most useful feature was the timelines, followed closely by the shortcut window, speed information in the data block, speed advisories, route modification tool, and color-coding of information. All of these features were rated as easy to use. The conflict list received the lowest mean usefulness rating and was neutral on the usability rating.

One reservation controllers expressed were with the route modification tool. The feeling was that the tool created too much clutter, particularly during high traffic times. One controller stated that it created too many lines for too long a period. Some controllers pointed out that they rarely used the conflict list, indicating that it did not draw one’s attention as readily as it should.

The TRACON controller indicated that the display was very easy to use and did not create any clutter problems during approach spacing procedures. He indicated that the additional self-spacing information in the data block, the expanding data block advisories, and the history circles were vital during these operations. He was neutral on the usefulness of the capability to click on FMS trajectories and mouse operations, and indicated that the shortcut window for handoff and runway change information was unnecessary.

4.6 SIMULATION ENVIRONMENT CONSIDERATIONS

The DAG-TM concept is evolving as research continues. Researchers continue to advance and adapt the simulation environment to meet the requirements for testing the DAG-TM concepts. The September 2002 DAG-TM simulation represented the current state of development at that time. We recognize that not all of the desired elements were in place. This section recounts several simulation environment issues, not directly related to the concepts that were raised during the 2 weeks.

4.6.1 Flight Deck

During the simulation, pilots and observers noted some problems with the PCPlane and ACFS displays and interface. Although there were issues specific to either the PCPlane or the ACFS, general consensus was reached on major issues.

Use of Tools

In general, several of the pilots' comments dealt with interacting with the RAT, MCP, and FMS. Many of these comments may be the result of the pilots' lack of experience with the specific tool; however, there were times when the interface did not appear to behave as expected.

- When initiating a route change through the RAT, pilots were required to press the EXECUTE button on both the RAT and the FMS. In this case, the extra step was required because the two systems were not integrated.
- There were cases where LNAV and VNAV did not behave as the pilots expected. VNAV would automatically, and inadvertently, disengage when the pilot entered altitude changes in the FMS. LNAV would not automatically disengage when the pilot entered a manual heading change, causing a new heading to be rejected.

Simulation Realism

A condition that pilots in both the ACFS and PCPlane conditions noted was that some of the aircraft capabilities, limitations and handling qualities were not realistic. For example, some of the speeds assigned by ATC were too fast or slow to accurately reflect current conditions and handling capabilities. Some pilots were concerned about stalling during some of the runs. Although these factors did not affect pilots' ability to interact with the system, they were a distraction.

Overall, pilots rated the communications with controllers as realistic. Pilots commented on the amount of chatter on the radios during the different conditions and different phases of flight. In current-day operations, a transmission typically occurs every 2 seconds in terminal airspace. In the simulations, the transmissions were much less frequent. The lack of chatter could be a simulation design issue or an indication of improvements in the system and procedures as a result of the DAG-TM concepts.

ACFS

Because the pilots in the ACFS were the only participants to perform as a coordinated team, some unique events were noted during the runs. There were occasional mismatches between the data on the two displays. For example, the two pilots would occasionally receive alerts at different times, or they would receive different alert levels for the same conflict.

Another problem in the ACFS was that old data link messages were not archived on the display the way they were in PCPlane. Sometimes the old message was replaced with a new message before the pilots had time to read the old one. In one instance, the message was displayed only for about 0.5 sec before being replaced by a new data link message. This incident emphasized that the pilots need some control over the display of the data link messages, such as a method to quickly retrieve or review the messages.

4.6.2 Air Traffic Control

Use of Tools

Controllers had some problems with TMA delay information – displayed near the traffic symbol and in the sequence list for the Baseline condition and in the timelines for the CE 6 and CE 5 conditions. When a flight was turned off its trajectory to absorb arrival delay, TMA could not accurately compute delay. Typically, the delay value would freeze and not update until the aircraft was turned back on its trajectory, at which time the delay value would recompute, generally showing an unexpectedly large change. Controllers found these relatively large variations in delay information confusing when trying to manage the schedule.

4.7 TRAINING

4.7.1 Pilots

Unlike the controllers, the pilots who participated in the simulation were all new to the DAG-TM experience. They were unfamiliar with DAG-TM, the concept elements, the tools, or the procedures. Training before the first test run consisted of a briefing pertaining to basics of CEs 5, 6, and 11 followed by a series of training runs. During the data collection runs, it became apparent to the observers that the training that the pilots received was not adequate for them to perform all of the functions necessary for effective task completion. Pilots seemed to struggle with some of the basic functionality of the CSD, FMS, and MCP. Pilots' ratings and comments on the post-simulation questionnaire also indicated that, for the most part, training on the PCPlane, procedures, and rules-of-the-road was not or only barely adequate. Because pilots experienced both concept-related and simulation-related difficulties, they may not have been able to differentiate between the two. As a result, their impressions of the actual DAG-TM concepts may have been skewed negatively by simulation-related difficulties; and system performance may have suffered.

Misinterpretation of Display Symbology

On the basis of pilot comments, it was determined that they did not always understand the significance of the display symbology. For example, one pilot thought that the shaded 200-mile circle around the airport indicated the freeze horizon. This shaded circle was actually designed to aid pilots in determining if the aircraft were within 200 miles of the airport so that they could apply rules-of-the-road with respect to arrival aircraft. The pilot's misunderstanding could have resulted in his or her not recognizing at which point en route aircraft become arrival aircraft; therefore, they would not properly apply the rules-of-the-road.

Procedures Misuse

For the most part, pilots were able to appropriately apply the general procedures for the three different conditions tested. However, pilots were not always clear on the specific procedures required for the current conditions. Some of these issues were related to procedures for interacting with the tools, such as the sequence of button presses required to load and execute an ATC reroute. In a few cases, pilots did not act appropriately for the condition being run. For example, in one case the flight crew sent a reroute request to ATSP during CE 5 operations. Whether this action was the result of the pilot's confusion about the procedures for CE 5, or whether the pilot was confused as to which condition was being run is not known.

Resolution Strategy

Air traffic controllers, who are currently responsible for maintaining separation among aircraft in the airspace, have had much experience in implementing strategies to resolve different conflict geometries. They can determine what resolution is most effective for most conflict geometries. Pilots, however, had never been responsible for resolving conflicts and therefore did not have the required experience in determining the most effective maneuvers. In some cases, pilots exerted significant effort trying to determine an appropriate resolution but were unable to find one.

4.7.2 Controllers

All controllers who participated in the September 2002 simulation had participated previously in DAG-TM events at ARC and were therefore familiar with the concepts, tools, and procedures. Controller training consisted of a review of tools and procedures that they had been exposed to before and training on any changes since they last participated in a DAG-TM event. The controllers rated the training they received as adequate. However, one controller commented that during training, experimenters should identify the concepts specifically and demonstrate the use of the concept so they are more easily recognizable when they are being tested.

5. CONCLUSION

In September 2002, AATT researchers conducted the first large-scale, high-fidelity human-in-the-loop simulation study involving air traffic controllers and carrier pilots using advanced technology and procedures developed for a distributed air/ground environment. The DAG-TM team succeeded in implementing an operational simulation environment encompassing multiple users and multiple locations. Pilots and controllers realistically accomplished their assigned tasks and even showed some system efficiency improvements in experimental conditions. The simulation clearly demonstrated that an air traffic management system involving a distribution, although limited, of the maneuvering and separation responsibility could be implemented.

The present study achieved many objectives, such as investigating the potential benefits and feasibility of DAG-TM CEs in a realistic and fully operational environment. The measurement of multiple system-level metrics and participants' reports will help in refining the concepts, proposed procedures, and ground and flight deck management tools used during the simulation.

5.1 BENEFITS

Results strongly suggest that the CEs, as implemented in the simulation, and the advanced technology used by controllers and pilots improved system performance. Inferential and descriptive analyses of arrival delivery, spacing, altitude deviation, en route altitude, distance traveled, and flight time system-level metrics all suggest that runs in the CE 5 and CE 6 conditions were more efficient than in the Baseline. Moreover, participating pilots and en route controllers rated the efficiency of the CEs favorably with respect to current-day operations. Controllers reported that maintaining arrival schedule was easier in the CE 5 and CE 6 conditions than in the Baseline.

5.2 FEASIBILITY

The feasibility of the CEs examined in the present study depends on factors such as the level of safety that they would provide and their acceptance by users. On the basis of analysis of system-level metrics, it appears that the gains in efficiency observed in the CE 5 and CE 6 conditions were not achieved at the expense of safety. More specifically, analyses indicate that meter fix crossing restrictions and minimum spacing requirements were met at least as often in the CE 5 and CE 6 conditions as in the Baseline. These results suggest, at least indirectly, that CE 5 and CE 6 did not increase the likelihood of traffic conflicts and therefore did not negatively impact safety.

Pilots indicated that they found that the level of safety offered by each CE comparable to or higher than in current-day operations. They also positively rated the acceptability of all CEs. Controllers' opinions regarding the safety and acceptability of the CEs were not as positive. Although they felt that they had been successful in maintaining safety in all simulation conditions, they found that CE 5 and CE 11 offered levels of safety below those of normal current-day operations. However, controllers found the level of safety provided by CE 6 comparable to normal current-day operations.

Another indication that the CEs did not negatively affect system safety is that a majority of the pilots and controllers responded that their workload remained at an acceptable level, or even slightly lower than during normal current-day operations. Controller reports and ATWIT ratings suggest that their workload remained the same or even decreased slightly in CE 5 and CE 6 conditions. The same pattern was reported for mental demand, effort, and frustration. Similarly, pilots estimated that their situation awareness was superior during the simulation compared with normal current-day operations.

Other results suggest that the DAG-TM CEs are feasible. For example, pilots indicated that the ease of use and safety of the simulation procedures and rules-of-the-road were comparable or superior to normal current-day operations. Similarly, controllers positively rated their general performance and reported monitoring and maintaining separation during the simulation runs with relative ease.

In summary, pilots considered that the CEs provided levels of efficiency and safety comparable or superior to those found during normal current-day operations; and they rated the acceptability of the CEs positively. Controllers rated CE 6 similarly to pilots and found that CE 5 and CE 11 provide potential efficiency benefits. However, they expressed reservations regarding the safety of CE 5 and CE 11, and did not rate the acceptability of the latter concepts positively.

6. LESSONS LEARNED

6.1 CONFLICT RESOLUTION

The rules-of-the-road, as they were implemented for this study, did not always enable pilots to make informed decisions about which aircraft was responsible for maneuvering. If the rules-of-the-road continue to be the only method of determining maneuvering responsibility, the rules must be further developed and tested to account for all conceivable conflict geometries, including multiple-aircraft conflicts and conflicts with airspace hazards. Similarly, the flight deck tools did not always provide the efficient means for resolving traffic conflicts. The task of generating a flight path modification became increasingly challenging when LOS was imminent.

A possible solution may be to research the feasibility of automated resolution advisories. These advisories would provide pilots with the best solution for a given conflict geometry. If an immediate resolution is not required, the "best" resolution may be determined by a number of factors, such as fuel efficiency or least impact on STA/RTA. These parameters could conceivably be preselected by the pilot so that the resolutions are somewhat customized per pilot or flight priorities.

6.2 ASSIGNED ARRIVAL TIME

The assigned arrival time (i.e., RTA, STA) was a powerful concept that was important in linking air and ground. TMA-generated RTA/STAs provided controllers and pilots with a common perspective on the traffic flow and likely facilitated air-ground coordination and system efficiency. The significance of this concept was well demonstrated by the desire of pilots to maintain their scheduled arrival time, as well as consider the RTAs of other aircraft when resolving traffic conflicts.

6.3 AIR-GROUND PROCEDURES

Observations made during the study suggest that air-ground procedures were not always followed or understood. At times, participants found that air-ground procedures were not clear enough or did not seem appropriate, or that they simply forgot them. In one situation, participants were not sure if autonomous control should be reinstated after having transferred control to the controller. This example emphasizes the need to establish unambiguous air-ground procedures. To determine their suitability in future studies, air-ground procedures will need to be formally established and sufficient training provided for participants to fully integrate them. Human-in-the-loop simulations will not be able to verify the suitability of proposed procedures if they are too difficult to learn or if participants do not have enough training.

6.4 FLIGHT CREW COORDINATION

Section 4.4.6, Flight Crew Coordination, describes an incident where the lack of awareness of the other pilots' actions and apparent lack of communication led to undesired consequences. This incident highlights a couple of issues. First, pilot-first officer coordination on the DAG-TM flight deck should be further investigated. Pilots need to coordinate their efforts so those tasks can be completed effectively and efficiently, and without negatively impacting the other's efforts. Second, specialized training likely will be necessary to develop the coordination required to operate in the DAG-TM environment. Future studies should focus more on flight crew coordination issues and, if possible, include more opportunities for pilots to work as crews rather than independent users.

6.5 CONCEPT TESTING

The purpose of the DAG-TM simulations was to study the feasibility of the tools and procedures that were developed to allow airspace users to function in a distributed airspace environment. A major thrust of this effort is on pilots' ability to free maneuver and negotiate trajectory changes. Trajectory negotiation involved many steps, and pilots found PCPlane difficult to use, which may have limited the investigation of the CE 6 tools and procedures. Traffic conflicts and airspace hazards are the main reasons for initiating a maneuver. If there are no conflicts to resolve or airspace hazards to avoid, pilots have little incentive to maneuver other than to meet their RTA, which often can be accomplished through speed changes. Several CE 5 and CE 6 runs were completed without the pilot ever having to initiate a maneuver. As such, the concepts were not tested to their fullest potential. Providing reasons for pilots to maneuver, whether to avoid a conflict with other traffic or to find a route around a weather cell, would increase the realism of the scenarios and allow for a more thorough testing of the operational concepts.

Similarly, the simulation included only a small number of autonomous aircraft in the CE 5 and CE 6 conditions. The DAG-TM concepts depict a future airspace that will include a much larger percentage of autonomous aircraft. Because there were relatively few autonomous aircraft,[3] the operational concepts may have not been fully assessed. An increase in the percentage of autonomous aircraft may be necessary to fully evaluate operational benefits such as safety and efficiency. Performance comparisons of equipped and unequipped aircraft in the present study were not possible due to several differences between the equipped PCPlanes and the unequipped MACS. For example, PCPlanes are single-aircraft flight deck stations, while MACS are multi-aircraft stations. Also, PCPlane flights, on average, started earlier in the scenarios and farther out from the meter fix than the MACS flights. To analyze equipped versus unequipped aircraft performance, at least some unequipped aircraft should be flown by participant pilots via a PCPlane flight deck station.

Future simulations and studies could include realistic airspace hazards, such as weather cells or special use airspace, and a representative mix of autonomous and managed aircraft. If the number of autonomous aircraft cannot be increased substantially, a possible solution may be to fly all autonomous aircraft in a concentrated area, which would increase the potential of interaction among autonomous aircraft.

6.6 PARTICIPANT TRAINING

Section 4.7, Training, listed examples in which participant pilots misunderstood the concepts, tools, and procedures that were implemented in the DAG-TM simulation. These examples indicate that training needs to convey information about even small nuances of the concept elements, symbology, procedures, and rules-of-the-road. New DAG-TM participants will require comprehensive training with a proficiency assessment before participating in the training runs to ensure that they have the knowledge and experience to perform proficiently. Advanced training may be necessary in some cases. For example, pilots may need to be trained on specific conflict resolution strategies that will resolve the conflict without taking the aircraft off its scheduled time.

---

[3] Approximately 8 percent of the total flights and 18 percent of DFW arrivals were under autonomous control in the CE 5 condition.

7. REFERENCES

Abbott, Terence S. (2002). *Speed Control Law For Precision Terminal Area In-Trail Self-Spacing* (NASA/TM-2002-211742). Hampton, VA.

Advanced Air Transportation Technologies (AATT) Project (1999). *Concept Definition for Distributed Air-Ground Traffic Management (DAG-TM) Version 1.0*. National Aeronautics and Space Administration, Aviation Systems Capacity Program.

Byrne, M.D. (1993). A better tool for the Cognitive Scientist's toolbox: Randomization statistics. *Proceedings of the Fifteenth Annual Conference of the Cognitive Science Society,* 289-293. Hillsdale, NJ: Lawrence Erlbaum.

Couluris, G.J. (2000). *Detailed Description for CE 6: En Route Trajectory Negotiation.* (Contractor Report NAS2-98005 RTO-41), Moffett Field, CA: NASA Ames Research Center.

Edgington, E.S. (1980). *Randomization tests*. New York: Marcel Dekker.

Federal Aviation Administration (1997). Federal Aviation Regulations: Part 91, general operating and flight rules. Washington, DC: Federal Aviation Administration.

Howell, D.C. (2001). *Resampling statistics: Randomization and the bootstrap*. Retrieved October 30, 2002, from http://www.uvm.edu/~dhowell/StatPages/Resampling/Resampling.html.

Noreen, E. (1989). *Computer-intensive methods for testing hypotheses*. New York: Wiley.

Phillips, C.T. (2000). *Detailed Description for CE 5: En route free maneuvering* (NAS2-98005 RTO-41). Moffett Field, CA: Technical Research in Advanced Air Transportation Technologies.

Prevot, T., Palmer, E., Smith, N., & Callantine, T. J., & (2002). *A Multi-Fidelity Simulation Environment for Human-in-the-Loop Studies of Distributed Air Ground Traffic Management*. American Institute of Aeronautics and Astronautics (AIAA).

Raytheon ATMSDI Team (2002). *DAG-TM Test Plan*. NASA Ames Research Center, Moffett Field, CA.

Raytheon ATMSDI Team (2002). *The NASA Ames Baseline Cockpit Situation Display.* NASA Ames Research Center, Moffett Field, CA.

RTCA. (1995). *Final report of RTCA task force 3: Free flight implementation*. RTCA, Washington, DC.

Sorensen, J.A. (2000). *Detailed description for CE 11. Terminal arrival: Self-spacing for merging and in-trail separation* (NAS2-98005 RTO-41). Moffett Field, CA: Technical Research in Advanced Air Transportation Technologies.

Stein, E.S. (1985). *Air traffic controller workload: An examination of workload probe* (DOT/FAA/CT-TN84/24). Atlantic City, NJ: DOT/FAA Technical Center.

APPENDIX A - ACRONYM LIST

AATT	Advanced Air Transportation Technologies
ACFS	Advanced Concepts Flight Simulator
ADRS	Aeronautical Data link and Radar Simulator
ADS-B	Automatic Dependent Surveillance-Broadcast
AOC	Airline Operations Center
AOL	Airspace Operations Lab
ARC	Ames Research Center
ARTCC	Air Route Traffic Control Center
ATA	Actual Time of Arrival
ATC	Air Traffic Control
ATM	Air Traffic Management
ATMSDI	Air Traffic Management System Development and Integration
ATSP	Air Traffic Service Provider
ATWIT	Air Traffic Workload Input Technique
CD&R	Conflict Detection and Resolution
CE	Concept Element
CNS	Communication, Navigation, and Surveillance
CPA	Closest Point of Approach
CPDLC	Controller Pilot Data Link Communication
CSD	Cockpit Situation Display
CTAS	Center TRACON Automation System
CVSRF	Crew Vehicle Systems Research Facility
DA	Descent Advisor
DAG-TM	Distributed Air/Ground Traffic Management
DAL	Dallas Love Field
DFW	Dallas/Forth Worth International Airport
DST	Decision Support Tool
ETA	Estimated Time of Arrival
FAF	Final Approach Fix
FANS	Future Air Navigation System
FD	Flight Deck
FL	Flight Level

FMS	Flight Management System
GA	General Aviation
LNAV	Lateral Navigation
LOS	Loss Of Separation
MACS	Multi Aircraft Control System
MCP	Mode Control Panel
NAS	National Airspace System
NASA	National Aeronautics and Space Administration
NM	Nautical Mile
PAS	Primary Alerting System
PC	Personal Computer
PFD	Primary Flight Display
PGUI	Plan-View Graphical User Interface
PTA	Proposed Time of Arrival
RAT	Route Analysis Tool
RTA	Required Time of Arrival
RW	Runway
SOP	Standard Operating Procedure
STA	Scheduled Time of Arrival
TCAS	Traffic Alert & Collision Avoidance System
TFM	Traffic Flow Management
TLX	Task Load Index
TMA	Traffic Management Advisor
TRACON	Terminal Radar Approach Control
WAK	Workload Assessment Keyboard
VCR	Videocassette Recorder
VNAV	Vertical Navigation
ZAB	Albuquerque ARTCC
ZFW	Forth Worth ARTCC
ZKC	Kansas City ARTCC

APPENDIX B - RANDOMIZATION STATISTICS

The purpose of this appendix is to familiarize the reader with the randomization test, a procedure for determining the statistical significance of hypothesis tests.[4] The following example, taken from Byrne (1993) will help illustrate this statistical test (see Siegel & Castellan, 1988; and Edgington, 1980, for similar examples), also called the permutation test. Consider the case of two small independent samples, where group A includes 4 subjects, and group B, 5. Each group includes the following observations:

Group A: 12, 7, 4, 3

Group B: 8, 10, 12, 15, 22

The null hypothesis is that the two samples are from the same distribution (or in other words, that their average are the same). The first step consists in choosing a test statistic that will measure the difference between the two groups. For our example, the difference between the two means will be appropriate. The difference between the means of group A (M=6.5) and B (M=13.4) is 6.9.

According to the null hypothesis, if random assignment was used (i.e., if subjects were randomly assigned the two groups), any rearrangement (or permutation) of the data is equally likely. The number of possible permutations in the present example is

$$\binom{n_A + n_B}{n_A} = \binom{9}{4} = \frac{9!}{5!4!} = 126.$$

The next step is to calculate the difference between the two means of each permutation, and count how many times it will be equal or superior to 6.9, the difference between the means of the observed data. In the present example, assuming a one-tailed test, there are 5 permutations, including the observed one, with a difference between the means equal to or larger than 6.9. The statistical significance (p) of the observed data is established by dividing 5 by 126 (p = 5/126 =.04). This means that if the null hypothesis were true, the probability to obtain the observed data would have been .04. If compared to a nominal alpha level (conventionally .05), the null hypothesis would be rejected, indicating that the two observed means were significantly different.

The present example employed the difference between the two means as a test statistic, but others could have been used, such as the difference between the totals of the two groups. All these test statistics would have led to the same results (Howell, 2001). Another equivalent test statistic is the t value. Using the t value as an index of difference between the two groups, the same method as in the example would be used. The t value of all 126 unique permutations would have been compared to the t value of the observed data (t=7.9), and the statistical significance of the test determined by the proportion of t values equal or superior to 7.9. Even if the randomization test uses a test statistic such as the t value, it remains a nonparametric (or distribution-free) test because it does not determine the significance of the t value based on the studentized significance table, but by using a distribution based on permutations of the empirical data.

[4] Complete descriptions of randomization statistics are included in Edgington (1995), Good (1994, 1999), Lunneborg (1999), Manly (1997), and Sprent (1998).

Although the concept of the randomization test was introduced more than 65 years ago (Fisher, 1935; Pitman, 1937a, 1937b, 1938), this method has gained in popularity only recently. One reason is the greater computational power available to researchers. The number of permutations augments rapidly with the number of data points. For example, a design with 3 groups of 9 cases per group has over 4 million permutations, and the same design with 10 cases per group has more than 30 million. An exact randomization test, which considers all possible permutations, may become unreasonably long to perform when the number of permutations is too large. Another option is to conduct an approximate randomization test, which uses only a randomly selected sub-sample of the permutations (usually a large number, such as 10,000).

The randomization test, applied to a two independent sample design in the previous example, is applicable to most experimental designs. Edgington (1980) extensively demonstrated how to analyze different experimental designs using typical parametric values such as t and F, but by determining their statistical significance directly from the data, without recourse to significance tables.

Many of the randomization tests performed in the present study were applied to a one-way repeated-measure design and used the F value as a test statistic. One example is the randomization test performed to determine if delivery error varied significantly across conditions. Table 11 presents the mean absolute arrival delivery error for all runs, according to experimental condition. The F value for the observed data was 18.6. To determine the significance (p value) of the test statistic, it was compared to a distribution of F values calculated from permutations of the observed data. To reduce the influence that a learning effect may have had on delivery performance, the analysis generated the rearrangements by permuting only the mean scores located on the same row in Table 11 (on the same run number). For example, the first baseline mean had to be permuted with the CE-6 or CE-5 score obtained during the first run, not during the second, third, or fourth runs. The number of possible permutations was therefore $6*6*6*6=1,296$. Because the software used in the study (Howell, 2001) could only perform approximate randomization tests (instead of exact tests), the observed data F value was compared to 100,000 randomly selected permutations. Out of these 100,000 generated permutations, 3705 had an equal or greater than value of 18.6, indicating that $p=3705/100,000=0.04$. Since p was smaller than .05, the nominal alpha level used in this report, this analysis suggests that absolute delivery error varied significantly between conditions.

Table 1. Average Absolute Arrival Delivery Error According to Condition and Run.

	Condition		
	Baseline	**CE 6**	**CE 5**
1st Run	68.3	14.7	15.0
2nd Run	47.0	19.8	17.3
3rd Run	81.2	14.7	18.3
4th Run	43.3	20.8	17.8

Some authors insist that parametric statistical tables are applicable only to random samples (Hays, 1972, as cited in Edgington, 1980), and random sampling is rarely employed in ATC simulation research, as in most behavioral science studies, and would be almost totally impractical. One advantage of the randomization is that it assumes only random assignment, not random sampling. Other advantages include that the test has no assumption regarding the distribution of the samples.

REFERENCES

Byrne, M.D. (1993). A better tool for the Cognitive Scientist's toolbox: Randomization statistics. *Proceedings of the Fifteenth Annual Conference of the Cognitive Science Society,* 289-293. Hillsdale, NJ: Lawrence Erlbaum.

Edgington, E.S. (1980). *Randomization tests*. New York: Marcel Dekker.

Edgington, E.S. (1995). *Randomization tests* (3rd edition). New York: Marcel Dekker.

Fisher, R.A. (1935). *The design of experiments*. Edinburgh: Oliver & Boyd.

Good, P. (1994). *Permutation tests: A practical guide to resampling methods for testing hypotheses*. New York: Springer-Verlag.

Good, P. (1999). *Resampling methods: A practical guide to data analysis*. Boston: Birkhaüser.

Howell, D.C. (2001). *Resampling statistics: Randomization and the bootstrap*. Retrieved October 30, 2002, from http://www.uvm.edu/~dhowell/StatPages/Resampling/Resampling.html.

Lunneborg, C.E. (1999). *Data analysis by resampling: Concepts and applications*. Pacific Grove, CA: Brooks-Cole.

Manly, B.F.J. (1997). *Randomization, bootstrap and Monte Carlo methods in biology* (2nd ed.). London: Chapman & Hall.

Pitman, E.J.G. (1937a). Significance tests which may be applied to samples from any populations. *Journal of the Royal Statistical Society Series B*, 4, 119-130.

Pitman, E.J.G. (1937b). Significance tests that may be applied to samples from any populations II: The correlation coefficient test. *Journal of the Royal Statistical Society Series B*, 4, 225-232.

Pitman, E.J.G. (1938). Significance tests that may be applied to samples from any populations III: The analysis of variance test. *Biometrika*, 29, 322-335.

Siegel, S., & Castellan, N.J., Jr. (1988). *Nonparametric statistics for the behavioral sciences* (2nd ed.). New York: McGraw-Hill.

Sprent, P. (1998). *Data driven statistical methods*. London: Chapman and Hall.